

Can there be an end to spying?

Public espionage, a prolific source of scandal this year, is a serious cause of international friction, a restraint on good relations and even a perversion of national civility. Can it be restrained?

SPIES and the public agencies employing them seem to be half way through another bad year. In the United States, a rash of prosecutions of ex-servicemen and ex-agents has made plain the uncomfortable if unsurprising truth that, in a market economy, even loyalty has a price. There has been particular chagrin in the United States at the revelation, unconvincingly disputed, that even the Israeli government has spent resources (from the US military aid budget?) keeping tabs on the government which is its own chief political and economic supporter. The United States has also found itself alternatively able to boast of a harvest of counter-intelligence information gathered from defectors and then embarrassed by the discovery that defectors can defect back again. This painful double and even triple dealing, unfamiliar in the United States, is well-known in Western Europe, where 1986 (so far) has seen yet another eastern intelligence network unmasked in West Germany, the usual crop of expulsions of Soviet diplomats from western capitals (followed by retaliatory expulsions from Moscow) and a general reinforcement of the sense that secrets are no longer safe, whatever they may be. Spies, and their employers, must be asking whether their game is still worth playing.

The issue is further sharpened by the latest twist in the long saga of double-espionage in Britain, a matter of public record since the defection of two Foreign Office officials to the Soviet Union in the 1950s, in circumstances suggesting a tip-off from a sympathizer in counter-intelligence. Who might that have been? The first candidate to come to light was Mr Kim Philby, off and on employed as a British spy, who was recognized (after his disappearance to Moscow in the 1960s) to have been a Soviet agent as well. A second is Sir Anthony Blunt, himself a one-time member of British intelligence, afterwards curator of the Royal family's art collection, who was discovered to have recruited Philby as a double agent and who was stripped of his knighthood a short while before his death. A third is Sir Christopher Hollis, who was for several years the head of one half of the British intelligence network, that called MI5, and who is now (after his death) accused of having been another double agent. One result is yet another bout of speculation about the motives of those who become spies, the supposedly special motives of those who become double spies and the efficiency of the services which employ them.

Moles

The latest most bizarre development concerns Hollis and related matters. In 1981, the British prime minister, Mrs Margaret Thatcher, said that the accusations that he was a double agent had been considered and that they could not be sustained. But, last week, lawyers representing the British government before an Australian court which is being asked to restrain publication of a book by a British ex-agent, said they were prepared to acknowledge, for the purposes of that hearing only, the truth of the allegations against Hollis. (The British courts have already given the government an injunction preventing two British newspapers from serializing the disputed book and even from reporting its allegations.) The stratagem will have the effect of concentrating the court's deliberations on the claim that publica-

tion would be a breach of confidentiality, but it is bound to sound like a version of that confusing lawyers doublespeak typified by the defence of a person against a murder charge which asserts that the defendant acted out of self-defence and was, in any case, elsewhere at the time.

Whatever happens in the Australian courts, the British government is bound to find itself under further pressure to tell more when the British Parliament reassembles in October. The position that the validity of the allegations is irrelevant to the trial is legitimate enough, but it will be interesting to see whether the British government's lawyers can persuade an Australian court that a breach of its contractual relationship with an ex-employee can damage the Australian national interest. Meanwhile, the government's critics might usefully use the interval to clarify their complaint: is it the narrow complaint that the government (retrospectively) is incompetent in having allowed its intelligence services to be riddled with double agents or that these days, there should be more intelligent ways of conducting these affairs?

Vanity

Double agents, whatever their motives, are an inescapable part of the spy business. Those who doubt that should read an account of the mismanagement of the British Special Operations Executive (SOE), a secret organization set up in German-occupied France in 1940, which has been contributed to *Intelligence Quarterly* (a London-based newsletter) by one James Rusbridger, himself an ex-agent. SOE and its doings has become part of the mock-heroic folklore of British exploits during the Second World War. The chief objective was to help organize indigenous resistance to the occupation, but SOE was put together hurriedly, amateurishly and in a way that guaranteed that German intelligence would know what was afoot. One especially damaging double agent, according to Rusbridger, was a Frenchman called Dericourt who was from September 1942 both SOE's Paris agent for the transport of people into and out of France and a regular confidant of the German counter-intelligence agencies. When members of the SOE network were being arrested by the score (usually to be executed), how was Dericourt allowed to carry on? Rusbridger's arresting explanation is that his immediate boss in London, parachuted into France with a radio operator to discover what had gone wrong with the network's security, was warned by Dericourt that he would be arrested at a planned rendezvous and so sent his operator to the assignment and to a predictable death. The guilt engendered ensured safe conduct for the French double agent in the damaging months that followed.

What this unheroic tale shows is that all kinds of venality can contribute to the recruitment of agents and double agents and to their survival. But that does not explain the rash of double agents with which the professional intelligence services have been afflicted in the past half-century. Why have so many talented people chosen this tortured and, for some, disastrous way of earning a living? Part of the explanation is that given by Mr Andrew Sinclair earlier this year (see William Cooper's review in *Nature* 322, 317; 1986); now-unfashionable intellectual



coterie (in this case, the Cambridge-based "society of the apostles") may engender that blend of vanity and arrogance that tempts people to believe they can mould the course of history by the secret exercise of their own wits. But in reality it is a shabby business, not merely damaging for those concerned when they are found out and for those whom they betray meanwhile, but for the conduct of normal relations between intellectual communities throughout the world.

The ubiquity of the spy has become an abomination. In many parts of the world, scientists and other intellectuals are habitually supposed to be potential agents when they travel abroad — and may be suspected of potential treason when they return. Who knows how much of the present fragmentation between East and West of the science enterprise derives from the brooding suspicion of the intelligence agencies, vulnerable though they are themselves to infiltration, of the movement of people from one place to another?

That is merely one of the reasons why the time has come to put intelligence-gathering on a proper footing. Another, perhaps even more urgent, is that the repeated discovery of espionage networks in different places is a constant source of sourness in international relations. That intelligence services from time to time embark deliberately on subversive and even criminal activities, and must otherwise operate in some degree outside the normal rules of public accountability, is especially offensive. That many advanced societies must now put up with the indignity of self-imposed self-surveillance by telephone tapping and other such means is another part of the price that everybody pays for the dubious luxury of sustaining espionage. What can possibly be done to restrain not so much the curiosity of governments about what others do (which is legitimate), but the means by which they seek to satisfy that curiosity?

The best hope of progress must be technical, based on the use of electronic means of monitoring and analysing other people's electronic messages. This is the chief function of the National Security Agency in the United States and the General Communications Headquarters (GCHQ) in Britain, which in principle offer a means of learning about the military capabilities of other governments without requiring that people elsewhere should be actively disloyal. Is it entirely out of court that the major espionage governments should agree among themselves to rely on devices such as these wherever possible? (The fact that even organizations such as GCHQ may be penetrated by spies, as has been spectacularly demonstrated in Britain, is probably not nearly as relevant as it seems; those with secrets to hide should be able to guess which are at risk from listening posts and should plan accordingly.) Economic and political secrets, also common objectives of intelligence agencies, are less easily accessible to electronic surveillance, but are much more often within reach of an intelligent reading of other people's newspapers than the intelligence community would like to think. The underlying fallacy of much of what the intelligence services get up to is the assumption that each marginal element of information is almost priceless. Especially when the price already being paid includes a large element of self-imposed illiberality, this cannot be the case.

But an agreement to restrain mutual curiosity even if it were a formal convention, would hardly be sufficient by itself. Why not go further, and amend the Vienna Convention that regulates the use made of diplomatic missions in such a way as to deprive diplomats found guilty of espionage of the immunity from prosecution which they enjoy at present? There would be obvious difficulties in getting governments to agree that their nationals should be tried under other national laws, which vary enormously both in what they require as evidence and in penalties they impose but it should not be beyond the wit of well-intended people to develop a system of international criminal law to meet the need for an effective stick to beat the spies. The carrot, of course, is the benefit of the civility that would follow for the rest of us. □

Chernobyl made plain

This is what the Soviet Union should say next week about the Chernobyl accident, and why.

NEXT week, there will be a gathering in Vienna, under the auspices of the International Atomic Energy Agency, to discuss the report prepared by the Soviet Union on the reasons for the accident at the Chernobyl reactor at the end of April. Only the sketchiest information about the Soviet document is as yet available (see opposite), and is certainly not sufficient in itself to satisfy legitimate curiosity elsewhere about the accident and its consequences. The chances are that even the full report will not do that: most official statements hurriedly put together leave unanswered questions that are best dealt with in the kinds of technical discussions arranged for next week, and in which the Soviet Union has agreed to take part. So far, so good.

That there is a legitimate curiosity in what happened at Chernobyl may seem to many in the Soviet Union to beg the question of the right by which people elsewhere inquire into strictly domestic matters, such as the arrangements for generating electricity. The simple answer, that the Chernobyl accident polluted the territory of other nations, is only part of the reason why some kind of explanation must be forthcoming, and by itself only requires from the Soviet Union an estimate of the amount and composition of the radioactivity released, an attempt to agree on estimates of the damage that may have been done with those affected and, if appropriate, an undertaking that there would be negotiations about compensation. Nothing has been heard, as yet, about formal demands for compensation, which would have to be raised through diplomatic channels by the governments believing themselves to have been injured, not at technical meetings such as that planned for next week. Meanwhile, the Soviet Union has implicitly acknowledged a wider legitimate interest in Chernobyl by its agreement to appear in Vienna.

A second explanation, but equally insufficient, for the Soviet decision to say more than the bare minimum is the need to head off future trouble with now-nervous neighbours. Long before the accident, there were protests about the siting of reactors near the borders between France and Belgium and between Czechoslovakia and Austria. So far, there has been relatively little complaint about Soviet reactor policy from the countries most at risk, those of the Warsaw Pact in Eastern Europe, but that may require special explanation. It is common prudence to take whatever steps are possible to anticipate and counter such anxieties. But the Soviet Union has gone further by agreeing to participate at Vienna and by doing some homework in advance.

What else can the world legitimately ask of the Soviet Union? Because even the Chernobyl accident might have been worse, people elsewhere have a right to be assured that the management of these complicated machines is conducted with the zealous regard for worked-out procedures which is the nuclear industry's standard boast on its own behalf. As things are shaping up, the Soviet Union plainly intends to meet this demand indirectly, by using an account of how laid-down procedures were not followed to suggest that the same need never happen again. But such an account will not suffice. Inevitably, and properly, questions will be raised at Vienna about the arrangements now in place for regulating the use of nuclear reactors. By what means are matters like this decided? Who are the people concerned? These will be painful questions for the Soviet Union to answer, but they cannot easily be ducked. Beyond them are the even larger issues in which external curiosity is legitimate: what can be gleaned over the years about the effects of radiation on people exposed to excessive doses of it? And what, in any case, is being done to safeguard or even monitor those at risk? There is no requirement in international law that one country must account to others for its care of its own people, but these exceptional circumstances require concessions even in that direction. □

Official Chernobyl report

Timetable for a reactor disaster begins to emerge

THE Soviet report on the Chernobyl reactor accident, distributed to the Vienna delegations to the International Atomic Energy Agency last week, is said by those who have read it to be strong on details of the reactor design but thin on details of how the accident happened. But the sequence of events leading to the accident, and more details of the "unauthorized experiment" said to have been under way on 25–26 April, have emerged.

For the first time since the accident, there are suggestions in the Soviet document that the RBMK reactor design is not fault-free. Thus the report suggests that existing reactors of this type should be modified by means of longer control rods, 2.4 per cent uranium enrichment (now 2.0 per cent) and extra reactor sensors even if these measures "degrade the economics" of the reactor type.

The report is intended as the basis for a discussion of the accident to be held at Vienna next week. The confusion attending its appearance stems from what appears to have been a Soviet requirement that the agency itself should distribute the report to national delegations only, but allowing national delegations to deal with the document (382 pages of Russian text) as they see fit.

Other information about the content of the report has appeared in Sweden and from the Atomic Industrial Forum of the United States which on Monday this week released a brief summary of the sequence of events leading to the accident. The forum said that "at first glance the report appears to be complete and frank".

According to a report in the Japanese newspapers *Asahi Shimbun* the "experiment" was intended to determine the amount of energy that might be recovered from a spinning electrical generating set and the coupled steam turbine after the steam input had been disconnected. According to this account, more than one assessment of this kind was planned, using the generating set at Chernobyl called Number 7 and for this reason, the reactor providing steam for the generator was not shut down. However, the system that would have caused the reactor to shut down of its own accord was disabled.

One of the difficulties in reconstructing the course of events at Chernobyl reactor Number 4 appears to have been that the data-recording system used for monitoring normal operations had been switched to the monitoring of the electrical test, with the result that capacity for the recording of reactor data had been preempted.

As a consequence, the sequence of events has had to be inferred from a mathematical model of the reactor.

The sequence of events began roughly 24 hours before the first release of radioactivity, at 01:00 local time on 25 April, when steps were taken to reduce the operating power of the reactor. Twelve hours later, at 13.05, the Number 7 turbine generator was disconnected from the system, but its electrical output switched to turbine Number 8. These are two generating sets handling the thermal outputs of the damaged reactor, Number 4 out of four completed reactors at the site.

The rate at which the speed of the Number 7 generating set would have declined would have been determined by the magnitude of the external electrical load. The emergency cooling system was disabled at 14.00 on 25 April, but the report says that the control room explicitly instructed that the reactor should continue to operate.

When instabilities appeared, two extra pumps were added to the water-cooling circuit some three and seven minutes respectively after 01.00 on 26 April, but the Soviet document says that the reactivity of the reactor had reached the point, at 01.22, at which the reactor should immediately have been shut down. In the

event, the reports say, the operators persisted with their experiments on the turbines, including the closing of the stop control valve of turbine Number 8 four seconds after 01.23 on 26 April.

Within half a minute, the accounts continue, the reactor began to generate more power, and the operatives sought to close it down by means of the control rods. But in the event, the full entry of the rods was prevented by an explosion. A witness is reported as saying that he heard two loud explosions at that time and saw a fireball scattering sparks on the roof.

The Soviet report outlines a series of errors in safety procedures of which the reactor staff were guilty, including the disablement of all but six or eight of the control rods with which the reactor is equipped, and the disablement of the power control system that would normally maintain the power output near that for which the machine was designed. The result is that control of the reactor would have been more difficult. The decision to disable the automatic shut-down system for the reactor would have made it necessary to use manual procedures for controlling the power output in case of an emergency.

One of the puzzling features of the report is the statement that, after the explosions, the temperature of the uranium fuel increased to between 1,600 and 1,800 °C, hot enough to volatilize some fission products but not enough to vaporize the refractory components of the fuel. But that step could be the one that apparently led to the further release of radioactivity on 29 April. □

Environmental research

Making the most of Chernobyl

BRITISH scientists are hoping that a £1 million package of research projects can be put together to seize the scientific opportunities presented by the Chernobyl nuclear accident. Radionuclides were released into the atmosphere on a scale that would never be contemplated experimentally, giving a unique chance to study their pathways in the environment.

An expert committee organized by the Natural Environment Research Council, the Coordinating Group on Environmental Radioactivity, hopes to include researchers from government departments, the universities, the Central Electricity Generating Board, British Nuclear Fuels and the Meteorological Office in an integrated programme. Much of the information gained would be of direct benefit for it would aid understanding of the consequences of other accidental releases of radioactive material.

Studies carried out so far already reveal a few surprises. Radionuclides are much "stickier" than expected; once deposited

on vegetation they are not quickly washed into the soil. Thunderstorms, triggered as warm, moist air containing the plume of fall-out, prove to be particularly effective in washing radionuclides out of the atmosphere (see p.690 of this issue). These topics will be studied further: others needing attention are the overall pattern of deposition, which is highly heterogeneous; the storage of radionuclides in snow and their release on melting; and whether the radionuclide pulse will confuse other studies on radionuclides in the environment.

Any programme would best be integrated with an overall European effort. Already attempts are under way for collaboration within the European Economic Community and it is possible that some money might be found from the European Commission. Otherwise national programmes will have to be tied together and funds for the British part must come from other programmes — there is little chance of extra money being granted in present political circumstances. **Alun Anderson**

Rocket launchers

Japan enters the space race

AFTER the successful launch of the H-I rocket last week, Japan's rocket experts are bursting with confidence. For now nothing seems to stand in the way of Japan entering the commercial satellite launching business in the 1990's — nothing, except perhaps a small group of tuna fishermen.

The two-stage H-I rocket blasted off from the National Space Development Agency (NASDA)'s space centre on Tanegashima island southwest of Kyushu at 5.45 am on 13 August. About 4 minutes later, the domestically-built cryogenic second-stage engine burst into life boosting the launcher into an elliptical orbit 1,500 km above the Pacific. Coasting over South America about 50 minutes later, the second-stage reignited for a brief 20-second burn that transferred the launcher to a circular orbit before release of first the Experimental Geodetic Satellite, later named "Ajisai" (Hydrangea), and then the Japan Amateur Satellite-1, christened "Fuji" (Wisteria).

News that radiowaves from Fuji had been received in Chile was greeted with shouts of "banzai" at the headquarters of the Japan Amateur Radio League whose members have invested 5 years and 120 million yen (£0.5 million) in building Japan's first ham radio satellite. It is only to be hoped that Japan's 720,000 ham enthusiasts do not overtax the little 50 kg satellite. Takeshi Saito, the league's amateur satellite committee chairman remembers with embarrassment an incident three years ago when Japanese hams jammed the frequency band for communication with the US space shuttle *Columbia* and then vented their frustration by trading abuses such as "bakayaro!" — meaning "you idiot".

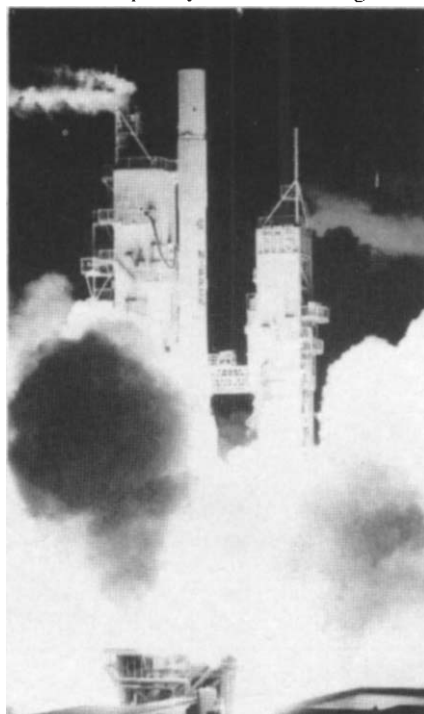
Equipped with 10 channels for analog communications and four for digital, Fuji comes over the Japanese archipelago six times a day permitting communications for about 10 minutes each time. A digital relay allows storage of messages which can be retransmitted on the other side of the Earth.

Ajisai, a 2-m diameter mirror ball to be used for triangulation surveys was also successfully put into orbit alongside Fuji. But the real hero of the day was the H-I rocket itself.

The LE5 second-stage cryogenic engine developed by NASDA at a cost of 44 billion yen (£192 million) performed flawlessly, as did the Japanese-built inertial guidance system. And in a spectacular finale five hours into orbit, the LE5 released a burst of liquid hydrogen fuel over the eastern United States, prompting reports of UFOs and exploding

satellites. This is the first time the US has found itself downrange of a nation with cryogenic launchers, and may be an omen of things to come.

Japan can now forge ahead with development of the bigger H-II which is scheduled for completion in 1991 and will be entirely made in Japan. The stage will then be set for Japan to enter the commercial market for satellite launchers. But will NASDA be able to increase launch frequency at its Tanegashima



Last week's launch utilized a US first stage but a Japanese cryogenic second stage.

Space Centre to meet the demands of foreign users?

The waters off Tanegashima teem with tuna, and under an agreement with the local fishing cooperative NASDA can only launch rockets during 90 days of the year, 45 days in summer and 45 in winter. This effectively limits NASDA to two launches a year, and for "compensation" NASDA annually pays the fishermen 600 million yen (£2.6 million). This is over and above the 1,300 million yen reputedly paid to the fishermen when the launch site was first built.

The only inconvenience to fishermen is that during the launch a 300 km area of ocean extending about 30 km downrange of the launch site is closed to shipping. But fishing cooperatives in Japan, like other rural groups loyal to the ruling Liberal Democratic Party, have considerable political power. And if NASDA wishes to increase launch frequency in the 1990s, the local fishermen will have the final say.

David Swinbanks

US space policy

No business for new shuttle

Washington

PRESIDENT Reagan last week ended months of uncertainty over the future of the US space programme by telling the National Aeronautics and Space Administration (NASA) to build another space shuttle orbiter to replace *Challenger*, thus restoring the size of the shuttle fleet to four. Reagan also signalled a major policy turnaround by saying that NASA would not in future be allowed to compete with the private sector to launch commercial satellites on the shuttle, but would instead concentrate on military and research payloads. The announcement marks the end of the policy of trying to make the shuttle commercially viable.

Senior administration officials have for some months been divided over how a replacement shuttle would be paid for, and last week's announcement from the White House did little to clarify where the required \$2,800 million will be found. The White House plans to ask Congress for \$2,350 million in extra budget authority for NASA over the next 5 years, using in part "unspent funds" from other government agencies. But it is unclear to what extent NASA's other programmes will be affected. Some additional belt-tightening at the agency seems a possibility, although Reagan renewed his commitment to build a space station.

The replacement shuttle will incorporate several improvements to the existing models. New brakes, engines and computer systems were mentioned, as well as other possible safety improvements recommended by the Rogers commission into the *Challenger* accident. The new orbiter is not expected to be completed until 1991.

Of the 44 commercial satellites NASA is already contracted to launch aboard the shuttle, only 16 will be launched by NASA before 1992. The remainder may be shifted back to commercial unmanned launch vehicles. But because it has been US policy to rely on the space shuttle as the nation's main launch system, US aerospace manufacturers have been unwilling to get into the private launching business, and even if there is now a scramble to manufacture new commercial launchers — which some doubt — many of the payloads that might have flown on the shuttle seem destined to be launched by foreign commercial services. Besides the French Ariane rockets, which are booked up for the next several years, both China and Japan have embryonic launching services that may take a significant proportion of the backlog of satellites now waiting to be launched.

Tim Beardsley

UK research

Call for more open councils

REFORM of Britain's research councils is called for in a new report from a body representing more than two-thirds of university teaching and research staff. The Association of University Teachers (AUT) wants to see more open research councils among a host of other changes it claims are necessary to deal with the "crisis" in university research.

The crisis is attributed to the breakdown of the dual-funding system. Although the universities carry out more than half of all basic research in the United Kingdom, the budget supplied through the University Grants Committee (UGC) (to maintain "well-founded" laboratories) and the five research councils (to support research projects) has declined in real terms while the costs of research are rising rapidly. Increasing numbers of researchers are joining the brain drain; libraries have had to cut back book purchases by 40 per cent since the 1970s; the student/staff ratio has deteriorated; and an increasing percentage of researchers face an uncertain life on short-term contracts.

The fight put up by UGC and the research councils to try to remedy matters seems not to have impressed AUT. UGC's new policy of trying to assess the quality of research carried out at different university departments and to concentrate scarce resources at those rated highly is seen as exacerbating the effects of the cuts and is "emphatically rejected". Selectivity, the report says, favours those areas in which short-term returns can be easily identified while university research should be of the "fundamental kind" from which "major breakthroughs rather than marginal advances result". The effect of selectivity may undermine university autonomy and create a hierarchy of universities.

The five research councils come in for some criticism. Their "success in producing growth areas in research has not been great", the report says. The councils are appointed by the Secretary of State but AUT believe they should be more directly representative of scientists. Nominations to the councils and their committees should be by learned societies and regular open meetings held to discuss priorities in research funding. Particular difficulties now arise because not all research proposals rated "alpha" (to be funded "at all costs") by peer review can now be funded. In that sense the report says "peer review has broken down" and personal views may be substituted. Safeguards to ensure that committee members do not receive preferential treatment, and full explanations of all decisions, including the reasons for choices between applicants, given publicly, are called for.

The Advisory Board for the Research

Councils (ABRC) could also be made more open, the report says. Each year, in its advice to government, the board has protested more strongly about the parlous condition of the universities. But the government shows little sign of listening. To add clout the report suggests that Parliament be given the right to discuss the advice of ABRC and to cross-question it. Whether this could be done without changing ABRC's status is not yet clear.

Creation science

Nobel laureates go into battle

Washington

SEVENTY-TWO US Nobel laureates have signed a brief filed with the Supreme Court this week urging the court to declare unconstitutional a Louisiana law requiring balanced treatment of evolution and creation in state schools. According to organizers of the brief, this is the largest group of Nobel laureates ever to sign a single document.

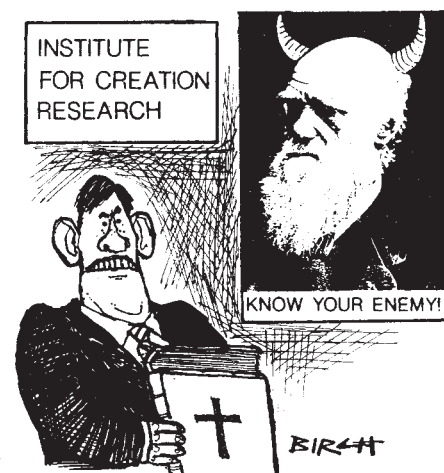
The Supreme Court agreed in May to hear arguments on the constitutionality of the Louisiana statute. In 1982, a federal district court struck down a similar Arkansas law because definitions in the law made it clear that creation science was directly derived from descriptions of creation in the Bible. The Louisiana statute removed the specific definitions of what creation science is, leaving only the tautology that creation science means "the scientific evidences [*sic*] for creation". The Nobel laureates' brief sets out to show that the Louisiana statute originally included language similar to the Arkansas statute, but lawmakers altered it after the Arkansas decision. The brief further attacks the law for singling out evolution as a theory, rather than emphasizing the uniformly tentative nature of science and argues that by creating a distinction between scientific



Murray Gell-Mann, one of the laureates leading the fight against the "forces of darkness".

An additional body is needed as a national forum to discuss research and development policy. The National Economic Development Council, an organization with representation from government, industry and unions, would be suitable.

British industry is seen as needing encouragement to commission research in the universities: there is evidence that foreign industrial investment in British universities exceeds that of British companies. Changes in tax-exemption rules and the establishment of a fund to help small firms sponsor university research are possible solutions. **Alun Anderson**



fact and evolutionary theory, the statute implicitly tells schoolchildren that "while most of what they learn in science class is 'proven scientific fact', evolution is not".

Murray Gell-Mann, who won the Nobel prize for physics in 1969, says it is not just biological science that has come under attack from creationists. Gell-Mann feels the Louisiana act represents a "grave danger to the Republic," one that will result in a decreased ability of schoolchildren to deal with scientific issues. Francisco Ayala, director of the Institute of Ecology at the University of California at Davis and a member of the National Academy of Sciences committee on science and creationism, agrees that nothing less than the "survival of rationality in this country's schools" is at stake. The Academy is preparing its own brief. Duana Gish, vice-president of the Institute for Creation Research (ICR) says his organization has not taken a position on the Supreme Court case.

The Supreme Court will hear oral arguments in the case sometime late this year or early next. Frank Press, president of the National Academy of Sciences says the debate should not be taken lightly. According to Gell-Mann, it is the "old struggle of science against the forces of darkness". **Joseph Palca**

Price-Anderson Act

Regulatory bill goes critical

Washington

AFTER months of hearings, meetings in smoke-filled rooms and political threats and counter-threats, Congress appears ready to extend the Price-Anderson Act. Due to expire next year, Price-Anderson establishes limits on the financial liability of commercial nuclear power plant owners in the event of an accident. With four different versions of the legislation currently floating around Congress, there still must be a lot of horse trading before a bill is passed. What finally emerges will go a long way towards determining the future of nuclear power in the United States.

The Price-Anderson Act, enacted into law in 1957, offered government indemnification to protect the fledgling nuclear power industry from financial losses resulting from a power plant accident. In its original form, Price-Anderson required utilities to insure themselves to the maximum amount commercially available — then \$60 million — and the government to provide indemnification for an additional \$500 million. In 1975, Congress modified the act, making utilities liable up to a fixed limit for claims not covered by their insurance. A \$5 million "retroactive premium" can be levied against utilities for each of their plants in the event of an accident. A key feature of the scheme is its "no-fault" provisions. Utilities will be liable collectively, regardless of which of them is at fault in an accident.

Today, with 101 operating plants and \$160 million in commercial insurance available, the total compensation to vic-

tims of a nuclear accident is \$665 million. The only real test of Price-Anderson was the accident at Three Mile Island, all claims related to which were settled out of the utility's commercial insurance cover.

A recent report* by the General Accounting Office (GAO) estimates that a \$665 million liability limit would be sufficient to settle claims related to a catastrophic accident for only 4 per cent of existing reactors. Even staunch supporters of Price-Anderson admit that the \$665 million figure is now unrealistically low. What the industry dreaded is that a new version of Price-Anderson would remove liability caps. But unlimited liability, vigorously supported by environmental groups, was shot down earlier this month when its principal congressional proponent, Senator Robert Stafford (Republican, Vermont) agreed to accept a realistically high, but fixed, limit on liability.

Current versions of the bills extending Price-Anderson vary widely in their liability limits. At one end of the spectrum, a bill sponsored by Representative Morris Udall (Democrat, Arizona) calls for retroactive premium payments of up to \$63-million per reactor per accident, limiting payments in a single year to \$10 million. The Udall bill also requires utilities to carry a minimum of \$200 million worth of insurance for each reactor, bringing the total liability limit for an accident to \$6,563 million. According to GAO calculations, this limit would be adequate to give full compensation for a catastrophic accident at 95 per cent of power

plants. At the other end of the spectrum, a bill reported from the Senate Committee on Energy and Natural Resources calls for a single payment, maximum \$20 million, per reactor per accident, for a total liability limit of \$2,180 million. GAO concludes that 64 per cent of reactors would be covered by this limit. The current prevailing view is that a cap in the region of \$5,400 will emerge in the final bill.

Industry reaction to the proposed changes in Price-Anderson is curiously mixed. Although industry may be forced to accept liability limits higher than it would like, the nuclear lobby appears to have prevailed on many aspects of its agenda. Environmental groups had sought changes in the law that would allow utilities to sue for recovery of their retroactive premiums from a negligent plant operator where an accident occurred. Keike Kehoe, director of the nuclear accountability and insurance project at the Environmental Policy Institute, says this would bring financial accountability to the industry, encouraging high safety standards to avoid huge financial losses. The industry position is likely to prevail.

Environmental groups are bitterly disappointed by the direction renewals are taking. Kehoe points out that in a 1983 report to Congress the Nuclear Regulatory Commission recommended shifting the full costs of a catastrophic nuclear accident to utilities, a position it has since backed away from. But Kehoe believes without unlimited liability, victims of a nuclear power plant accident may never receive fair compensation. Kehoe says the industry can provide \$10 million per year for as long as it takes to settle all claims. She points out that the industry is already committed to spending around \$5 million a year in insurance for each plant.

One contentious provision of proposed Price-Anderson extensions that could ultimately torpedo the bill's chances is indemnification of the government's own nuclear activities. Under current law, non-defence government nuclear activities have a \$500 million limit on liability. In some versions of the new legislation these liability limits would be removed entirely, both for nuclear plant activities and for nuclear waste activities.

If Price-Anderson is allowed to expire, all currently licensed nuclear plants would continue to be covered under existing law, meaning that liability would only rise slowly from its current \$665 million level as plants under construction come on-line. But without an extension to Price-Anderson, the next generation of power plants will face the prospect of unlimited liability. The industry says that means they simply would never be built. **Joseph Palca**

Change of priorities in French science

SUPPORT for basic research in France will be little changed next year, in real terms, from 1985, the last full year of the previous socialist administration, according to unofficial figures published by the newspaper *Le Monde*. But direct government aid for industrial research will be cut.

The level of support comes nowhere near that foreseen in the three-year research plan for which the French National Assembly voted last December. That would have given by 1987 some 8 per cent real growth in research budgets, and 2,800 new posts for scientists and technicians. The *Le Monde* report foresees just 280 new research posts and 500 technicians, engineers and administrators are to be sacked or not replaced.

The special funds of the ministry of research (spent directly in industry or government and university laboratories, rather than through research councils on applied topics) will also fall from FF1,200 million (£120 million) in 1985 to FF750

million in 1987. The new government is apparently not convinced that ministry research judgements, at least on the previous scale, are better than industry's own. ANVAR, the government agency promoting research in small and medium-sized industries, including the thousands of companies making up France's agro-food sector, can also expect to lose 30 per cent of its funds.

Senior government advisers appointed by the previous administration and still in post, do not think the new government has abandoned industry, however. They expect new tax breaks on research and development budgets to be announced. But many doubt whether such fiscal incentives will do much to change the present extreme concentration of French industrial research in a few major companies, or to treat the basic French malaise: poor links between academics and industry caused by traditional divisions in the education system. **Robert Walgate**

*Nuclear Regulation: Financial Consequences of a Nuclear Power Plant Accident (General Accounting Office, Washington DC, GAO/RCED-86-193BR).

Israel's water

Down to the very last drop

Rehovot

DESPITE having the most comprehensive water-planning system in the world and a plethora of internationally recognized water experts, Israel is now suffering the worst water shortage in its history. Government officials blame three years of drought, but most scientists blame the officials.

Dr Gerald Stanhill, an agricultural meteorologist at the Ministry of Agriculture's Volcani Centre, says that, statistically, the past three years were not exceptionally dry. Indeed, they fell well within the rainfall values to be expected from the pattern of the past 55 years. Over that period, annual rainfall within the geographical boundaries of Mandatory Palestine (the area controlled by Israel today) averaged eight cubic kilometres, with a standard deviation of 25 per cent. Perhaps, as the cloud seeders claim, seeding operations have somewhat increased rainfall since the late 1960s. However, the statistics certainly show that rainfall has not declined.

"Israel's water crisis is like its money crisis. It has simply drawn more out of its account than has gone into it", Stanhill says. He welcomes the government's decision to cut back water consumption by 10 per cent this year, although he thinks that what is needed is a policy to build up and maintain a national water reserve at a level equal to three years' consumption.

Dr Joseph Shalhevet, acting director of the Volcani Centre and chief scientist at the Ministry of Agriculture, is less alarmed by developments than other scientists. In his view, "the situation is tight but not impossible". If next winter is also dry, however, drastic cutbacks will be required, as they were, he points out, in California seven years ago. In that case, Israel, for some years a net exporter of agricultural products, would become an importer.

Meanwhile, Volcani scientists who are largely responsible for the exceptionally efficient use of water in Israel's agriculture, are looking for new ways of getting more dollars out of every drop. They have successfully introduced mangoes, avocados, spices, medicinal plants and new varieties of flowers. But each has only a limited export market, and all of them together cannot possibly replace water-hungry cotton, now the main cash crop of Israeli farmers.

Nevertheless, cotton acreage is being reduced this year, as an emergency measure by one-third, and some water experts — who point out that it has been grown at a loss in recent seasons — say it should be eliminated completely, even if this clashes with the Zionist dream and

drives thousands of farmers off the land.

Agriculturalists and their supporters are unwilling to write off cotton, pointing out that flowers, export earnings from which went up this year from \$74 million to \$104 million, were cultivated at a loss for some time and have now bounced back.

Advocates of cotton point out that it can be irrigated with sewage water, but this

could be potentially dangerous, says Weizmann Institute Professor Mordechai Magaritz and Daniel Ronen of the Israel Hydrological Service. Irrigation with treated sewage introduces undesirable amounts of pollutants into the soil, and unless comprehensive preventive measures are taken, into the aquifer as well. Magaritz and Ronen would therefore limit the irrigation of cotton with treated sewage to areas such as the semi-arid Negev, where there is practically no danger of polluting important aquifers.

Nechemia Meyers

Fossil bird shakes evolutionary hypotheses

Washington

FOSSIL remains claimed to be of two crow-sized birds 75 million years older than *Archaeopteryx* have been found in the 225-million-year old Dockum Formation near Post, Texas. Sankar Chatterjee, a palaeontologist at Texas Tech University, who found the fossils, says they have advanced avian features that place them closer to the ancestor of modern birds than *Archaeopteryx* and make them possible direct ancestors.

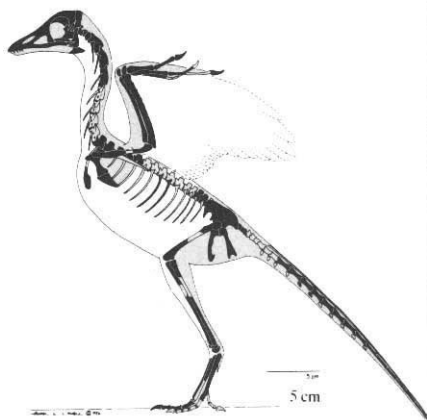
Chatterjee intends to assign the find to a new genus, *Protoavis*. Among the features that Chatterjee claims identify *Protoavis* as a bird are elongated forelimbs, temporal fossae confluent with the eye socket, a definite furcula (wishbone), and a plate-like sternum (breastbone) with a small keel. The quadrate is also said to be typically avian. Other avian features are a cnemial crest on the tibia and flight-muscle attachment sites on the humerus. No feather impressions have been found, but Chatterjee claims to see quill nodes on the ulna and hand bones and he believes that *Protoavis* could fly. *Protoavis* also has reptilian features: four teeth in the forward part of its jaw, a tail and clawed fingers.

Protoavis seems certain to reopen a long-running controversy on the evolution of birds, in particular whether the common ancestor of birds and dinosaurs was itself a

dinosaur. *Protoavis*, from the late Triassic, appears at the time of the earliest dinosaurs, and if the identification is upheld it seems likely that it will be used to argue against the view of John Ostrom of Yale University that birds are descended from the dinosaurs. It also tends to confirm what many palaeontologists have long suspected, that *Archaeopteryx* is not on the direct line to modern birds. It is in some ways more reptilian than *Protoavis*, and the period between the late Jurassic *Archaeopteryx* and the worldwide radiation of birds in the Cretaceous has to some seemed suspiciously brief.

Chatterjee's research has for the past six years been supported by the National Geographic Society, which publicized the find last week. The work has yet to be published in a peer-reviewed journal, but two independent palaeontologists (including Ostrom) who have examined *Protoavis* agree it should probably be classified as a bird. Ostrom stresses, however, that the remains are very fragmentary, and while agreeing with Chatterjee's tentative classification says the case is not finally proven. But Ostrom disagrees with Chatterjee in not seeing even indirect evidence of feathers, saying the bones where quill nodes might be seen are badly preserved, nor does he consider the forelimbs especially elongated.

Tim Beardsley



The fossil bones and an artist's reconstruction showing the recovered fossils (dark bones) of the putative *Protoavis* "bird".

Weapons potential

Defence lobby eyes antimatter

THE US Air Force appears to be taking seriously a report from the Rand Corporation that antimatter might be used as a component of weapons systems. Antimatter packs a tremendous punch, equivalent to some 44 tons of TNT per milligram, when it annihilates with ordinary matter, but until particle physicists began to make and handle antimatter (particularly antiprotons) routinely, any thought of using the material was mere science fiction. But the report, commissioned from Rand, a non-profit US think-tank, by the US Air Force in 1983 and released last year, claims that problems facing the use of antimatter in rocket motors for reversing direction in Earth orbit, beam weapons and the pumping of X-ray lasers could be resolved within five years. Sufficient amounts of antimatter to be militarily "interesting" — some 10 mg a year — could be produced within another 15 years if the US government were prepared to spend between \$5,000 and 15,000 million on a development programme.

Basic research on antiproton trapping is already under way at laboratories such as the European Organization for Nuclear Research (CERN) near Geneva. The report says that interaction with basic science could push a military programme for-

ward: "Our position is that the critical experimental work... is so rich with interest and so widespread in the area it intersects that researchers outside existing defense research (as well as those in it) should find it stimulating and an opportunity for creative invention". This interest leads to Rand's "reasonably confident prediction" that five years of research would resolve whether military applications are possible.

But could antimatter, currently produced in only minuscule amounts (some 10^{11} antiprotons per day, or 10^{-8} – 10^{-7} mg per year at CERN, for example) really be collected and used? According to Rand, the production and storage problems will be difficult and complex, and "there can be no confident assertion" that they will be solved. Many handling issues might be solved in experiments on normal matter (such as negative hydrogen ions that simulate the mass and charge of antiprotons) and others with present antiproton production technologies. Antiprotons could be held in static or dynamic electromagnetic traps (for example, Penning traps); at special sites in normal matter (for example in liquid helium bubbles); as spin-polarized antihydrogen; or as condensed antihydrogen in a cryogenic enclosure.

Antimatter has already been stored in minute quantities, and an experiment approved at CERN by a collaboration led by a Los Alamos group could lead to the storage of 10^{11} antiprotons per cubic centimetre. The main aim of this experiment is the collection and cooling of single antiprotons to measure their weight (do they fall upwards?), but the funding for this part of the experiment has not yet been raised. The rest of the experiment involves antiproton accumulation in a Penning bulk ion storage trap. The experiment has nothing to do with Rand's proposals, a spokesman points out.

Production of sufficient antiprotons to be useful, and the energy costs involved, may be the hardest problem facing weapons designers. Produced by colliding a proton beam to a target, at most one antiproton is collected per 100 incident protons, and this at the highest beam energy in use for the purpose (120 GeV at the Fermi National Accelerator Laboratory near Chicago). Taking into account the energy efficiency of accelerators, and assuming a "difficult near-term improvement" in antiproton collection technology, some 4 GW of external power would be needed to produce the militarily useful amounts of 10 mg of antiprotons a year, the report estimates. This is a huge and expensive power consumption but the gaseous diffusion plants for the separation of uranium isotopes once used in the manufacture of nuclear weapons consumed 6 GW. With the same parameters, antimatter production costs would reach \$133 million per mg. "Some applications" could easily accommodate such costs, the report says.

Antimatter could be used for propulsion, power generation, beam weapons and "classified additional special weapons roles". In propulsion, for example, the annihilation of antihydrogen with hydrogen could produce effective exhaust velocities of 10 km s^{-1} to a major fraction of the speed of light. Propulsion missions otherwise more or less impossible because of weight of fuel (such as Earth orbit reversal), would become feasible if antimatter were used.

In power generation, antimatter could be used for orbital prime power "for engagement". Antimatter could also help in the design of lightweight beam weapon systems, be used as beams for "hard kills", or help produce the very short energy pulses required in pumping X-ray lasers.

Understanding antimatter would be a "prudent" goal for the US Air Force and worth an investment of \$1 million on basic research as a preliminary study, and if that went well another \$100 million spent on a phase-A (pre-prototype production factory) study. The Air Force appears to have taken this recommendation seriously, and a research programme has begun.

Robert Walgate

Computers thriving from NSF coffers

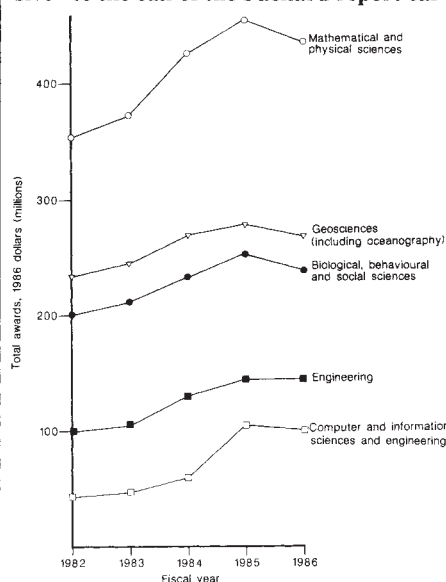
Washington

THE total value of research awards by the National Science Foundation (NSF) has increased substantially over the past four years, although some subjects have done much better than others (see below). The graph shows total awards made by each of the foundation's research directorates in constant dollars; retroactive adjustments have been made as necessary to take account of changes in jurisdiction, so that categories are consistent. Average award size by each directorate also increased over the same period.

All directorates' awards increased between 1982 and 1986, but a fall between 1985 and 1986 is evident, caused in part by the effect of Gramm-Rudman budget cuts in the current fiscal year. Much the largest increase was for computer and information sciences and engineering (129 per cent after inflation over the four years). This largely reflects NSF's establishment of five supercomputer centres. The other increases were 45.3 per cent for engineering, 23 per cent for mathematics and physical sciences, 14.4 per cent for geosciences and 21 per cent for biological and social sciences.

The administration proposed an increase again in 1987 for NSF, but Congress may yet balk at the idea (see *Nature* 322,

487; 1986). NSF's budget proposal for 1988, now ready to go the Office of Management and Budget, will, according to director Erich Bloch, be "highly responsive" to the call of the Packard report ear-



lier this year (*Nature* 321, 372; 1986) for NSF to administer a \$10,000 million 10-year programme to improve university research facilities.

Tim Beardsley

Penguin deaths questioned

SIR—Reports that the penguins in the Falkland Islands are dying on a large scale and that this may be due to competition for food from fishermen or marine pollution seem premature. I happened to be watching birds at sea off the Falklands during the season of the recent rock-hopper penguin (*Eudyptes chrysocome*) mortality in both 1985 and 1986.

Most of the Falkland seabirds appear to benefit greatly from the activities of the fishermen, following their boats in hundreds of thousands to feed on the spilt fish and offal from deep-water species that would not otherwise be available to them, and their populations are obviously flourishing.

The penguins are usually difficult to see at sea, but although they frequent the same area along the edge of the continental shelf they do not normally appear to follow the fishing fleets, either because, being unable to fly, they have difficulty in keeping up, or because they appear to feed in roughly equal proportions upon plankton and smaller squid of the genus *Teuthowenia*¹ which the fishermen do not appear to catch yet, taking larger fish and

squid of the genera *Illex* and *Loligo* instead.

The main difference between the two years in fact appeared to be that the southern summer of 1986 was unusually hot in the South Atlantic, with the result that among other things there was an influx of warm-water seabirds from the north, and the spring fires intended to remove the dead grass in the Falklands escaped control and burnt in the peat for months over areas of up to 6,000 acres, which one would have thought ought to be of much more concern to local naturalists. More penguins fed inshore than in 1985, possibly because the warm weather affected the growth or accessibility of their food at sea in the way that occurs when the warm current El Niño extends southwards off Peru in the Pacific.

A few dozen penguin bodies were seen both at sea and on the beaches, with more reported at the breeding colonies, but the number was not large by other standards⁶ for a population estimated at five million breeding birds⁷.

The main cause for concern appears to be that the huge breeding population on

Campbell Island, formerly the largest in the South Pacific and possibly the world, has already been declining on a much larger scale for over forty years⁸, though fortunately this appears to be offset by an increase on the islands off Tierra del Fuego (G.S. Clark *et al.* in preparation.).

Thus it would appear that the numbers of southern seabirds as well as those breeding in Britain⁹ have been fluctuating for longer and on a larger scale than there have been local overfishing and pollution problems.

W.R.P. BOURNE

3 Contlaw Place,
Milltimber,
Aberdeen AB1 0DS, UK

1. Lyster, S. *Falkland Island Fdn. Newsl.* 5, 2–4 (1986).
2. *Nature* 322, 4 (1986).
3. Croxall, J.P., Prince, P.A., Baird, A. & Ward, P.J. *Zool., Lond.* (A) 206, 485–496 (1985).
4. Beddington, J.R., Brault, S. & Gulland, J. *The Fisheries around the Falklands* (IIED/UCN Marine Resources Assessment Group, Centre for Environmental Technology, Imperial College, London).
5. Schreiber, R.W. & Schreiber, E.A. *Science* 225, 713–716 (1984).
6. Bourne, W.R.P. in *Marine Pollution* (ed. Johnston, R.) 403–502 (Academic, London, 1976).
7. Croxall, J.P., McInnes, S.J. & Prince, P.A. *ICBP Tech. Publ.* 2, 271–291 (1984).
8. Moors, P.J. *Polar Rec.* 23, 69–73 (1986).
9. Bourne, W.R.P. in *Enjoying Ornithology* (ed. Hickling, R.) 226–231 (Poyser, Calton, 1983).

Turkey maligned

SIR—Nechemia Meyers (*Nature* 321, 801; 1986) claims that visas had been denied by the Turkish Foreign Office to Israeli scientists seeking to take part in the Subcommission on Triassic Stratigraphy's field workshop (of whose organizing committee we are members) in Western Turkey between 14 and 23 July 1986. It seems, however, that Meyers was not particularly well informed, for Dr Mordeckai Magaritz of the Weizmann Institute, to whom a visa had allegedly been denied, had left Israel before a visa could be issued to him. It was subsequently requested that his visa be sent to him in Milan, Italy, where he was attending the meeting on the Permian and Permo-Triassic boundary in western Tethys held in Brescia, Italy, between 4 and 12 July. The Turkish Foreign Office obliged him, thanks to the timely intervention of Dr Remzi Akkök, but Magaritz then claimed that it was too late for him, although it was certainly not too late to attend the meeting in Turkey. Two other Israeli scientists, Dr Baruch Derin and Dr Tuvia Weissbrod, not only obtained their visas, but also attended the Turkish field workshop. One thus wonders whether there really was any difficulty associated with the visa that prevented Magaritz from coming to Turkey.

Turkey is indeed cautious about issuing visas to foreigners for scientific work and visits. The reason for this is a long history of scientific plunder, first made famous by Heinrich Schliemann, mainly in such

fields as archaeology, epigraphy and mineralogy, to which this country has been subjected. This has harmed not only Turkey but also science itself. No particular nation has therefore been immune to close scrutiny in this regard by the Turkish Foreign Office and there is no reason why Israelis should be an exception.

AYMON BAUD

Musée de Geologie,
Lausanne,
Switzerland

JEAN MARCOUX

Université Paris VII,
Laboratoire de Tectonique,
2 Place Jussieu,
Paris, France

A.M.C. SENGOR

İ.T.Ü. Maden Fakültesi,
Jeoloji Bölümü, Teşvikiye,
İstanbul, Turkey

African lakes

SIR—You published a year ago (*Nature* 315, 19; 1985) an article by Barel *et al.* entitled "Destruction of fisheries in African lakes".

Since then, a number of articles and reports have been published suggesting that Malawi plans to introduce the clupeid fish *Stolothrissa tanganicae* and *Limnothrissa miodon* (Kapenta or Ndaga), which occur naturally in Lake Tanganyika, into Lake Malawi (formerly Lake Nyasa).

The government of Malawi therefore wishes to make it clear that it has no current plans, nor has it had any plans in

the past, to introduce any exotic fish species into Lake Malawi.

Although the Malawi government wishes to maximize the sustainable fish yield from all its waters, it does not contemplate achieving this by resorting to the hasty introduction of exotics.

The views of the Malawi government on the subject of introducing clupeid planktivores into Lake Malawi are as follows:

- (1) There is at present insufficient knowledge of the chemistry, primary and secondary productivity, energy budget or trophic interactions of Lake Malawi to allow reliable predictions to be made about the outcome of any introduction.
- (2) Without such information, no introduction of any kind is justifiable.
- (3) If scientific studies suggest that the introduction of exotic clupeids could substantially increase the annual fish yield from Lake Malawi on a sustainable basis, without threatening the extinction of the lake's endemic fishes, Malawi would recommend the convening of multinational meetings, to include representatives of the three countries surrounding Lake Malawi and prominent fisheries scientists and biologists, to discuss and debate the issue before any action was taken.

G.M. NONGWA

(Acting Chief Fisheries Officer)

Ministry of Forestry
and Natural Resources,
Fisheries Headquarters,
PO Box 593,
Lilongwe, Malawi

Investigating the paranormal

SIR—In his Commentary on "Investigating the Paranormal" (*Nature* 320, 119; 1986) David Marks raises important issues about the problems of studying claims in areas where strong beliefs may be involved. With regard to parapsychology, however, his analysis (and that of many others quoted by him) appears to suffer from a lack of knowledge of the more process-oriented, systematic research that characterizes much present-day parapsychology. He claims that parapsychology is functioning without any theoretical bases, without any idea of what variables are relevant, in such a way that "every new investigator must start afresh, as though he or she is the first worker in the field". Yet much research in parapsychology explores the effect of experimental manipulations, as independent variables, upon measures of *psi* scoring as dependent variables. To appreciate this, one must be familiar with the details of original research reports. There are areas of research with partial replicability, that cannot yet be dismissed as "abortive leads"; current research does attempt to build upon them, as is evident from reading the introductory and discussion sections of research reports.

In his surveying of the value of the experimental research, Marks has relied upon two surveys of portions of that work, one by Akers and one by Hyman². Both surveys conclude that parapsychological studies have many methodological flaws. I agree that much of the work has been flawed, but there has been considerable debate about the extent of such flaws. Marks, for instance, has ignored a forty-page response to Hyman's article by Honorton³ which appeared in the same journal issue. The debate over specifics continues, and doubtless will for some time.

As Marks notes in his article, both John Beloff and I acknowledge that there has never been a perfect parapsychological experiment. What he does not explain, however, is that we argue that in areas studying extremely complex systems, as in the social sciences, it may be impossible to design, conduct and report a "perfect" study. In my article, I list several groups of counterhypotheses to any given study which, by their very nature, are difficult if not impossible to falsify, such as experimenter fraud or improper description of procedure. Rather, I argued, it is better to evaluate parapsychology by focusing on groups of studies that appear to find consistent relationships between a measure of psychic functioning and some other variable; that such studies would provide more relevant evidence than studies merely finding simple deviations from baseline.

I did not advocate examining "groups of studies which, although individually flawed reveal the undeniable presence of *psi*", nor did I advocate considering badly controlled experiments. As near as I can tell, Marks and I would agree that science progresses in its understanding by exploring functional relationships rather than merely documenting anomaly.

Lest I be misconstrued, certainly I agree that much of what has been done in parapsychology leaves something to be desired. Mistakes have been made, strong claims have been made by enthusiastic advocates, and theory construction has often been vague. Marks makes many good points about how advocates can be misled. But it is a mistake to assume that all who are willing to adopt the working hypothesis that there are new means of exchange between organism and environment are committed to a metaphysical belief system of some sort. If we are to improve our understanding of the phenomena explored by parapsychology, it will probably be through a more integrative approach, with more thorough evaluation of past work and better design of new research. An important component of such an integrative parapsychology would be the development of far more effective models for the strategies by which observers throughout the continuum of advocacy can be deceived by themselves and others about the nature of what they observe, perhaps along lines currently being developed by this writer⁵.

ROBERT L. MORRIS

2 Strathalmond Green,
Edinburgh EH4 8AQ, UK

1. Akers, C. in *Advances in Parapsychological Research* Vol. 4 (ed. Krippner, S.) 112–164 (McFarland, Jefferson, North Carolina, 1984).
2. Hyman, R. J. *Parapsychol.* 49, 3–49 (1985).
3. Honorton, C. J. *Parapsychol.* 49, 51–91 (1985).
4. Morris, R. L. *J. Am. Soc. psychical Res.* 74, 425–443 (1980).
5. Morris, R. L. in *Foundations of Parapsychology* (eds. Edge, H., Morris, R., Palmer, J., & Rush, J.) 70–110 (Routledge & Kegan Paul, London, 1986).

SIR—I should like to expose Marks's misleading simplification in the section of his paper on the paranormal headed "Coincidences".

Marks refers to Arthur Koestler² and L. E. Rhine³ as publishing cases easily subsumed under the heading of coincidence. However, Koestler, for all his services to psychical research, did not investigate cases and neither, strictly speaking, did Rhine, who simply scanned for recurring features in reports that correspondents sent her. With the exception of one of Rhine's cases, those published by her and Koestler may indeed be interpreted as "simple coincidences". Other cases cannot be so readily explained.

The Society for Psychical Research gave attention to possibilities of coincidence from its earliest studies of apparitions in the late nineteenth century. This explanation may be appropriate for many ex-

periences that laymen attribute to a paranormal process. They sometimes claim that a dream or hallucination of a fairly ordinary type has a paranormal link with a contemporaneous or future event that corresponds in a general way with the dream or hallucination. However, the events concerned in other cases are not ordinary, recurring ones, but unique events. A man dies only once in his lifetime, and if another person at a distance becomes aware of unusual details in his manner of dying at the time it occurs, something more than coincidence must be in play.

Consider the case of Mrs Agnes Paquet, who had a vision of her brother falling overboard when his foot got caught in a tow-rope on a tugboat (almost) at the time the brother actually drowned in this manner. Her vision included details of her brother's clothing, such as that his pants legs were rolled up so that, as he went overboard, she could "see" their white lining. She was approximately 50 miles away when she had her vision of the accident to which it corresponded. She told her husband about her vision before learning of her brother's death by telegram, and he corroborated this. The event was unique and so was the perception¹.

Students of paranormal phenomena, from Dr Johnson in the eighteenth century⁷ to Henri Bergson in the twentieth have emphasized that when details of unique events are correctly communicated over long distances without the normal sensory channels we should not dismiss such experiences as coincidences.

Many cases, perhaps hundreds, like Mrs Paquet's have been published. Why should we not continue to study them?

IAN STEVENSON

Department of Behavioral Medicine and
Psychiatry,
Box 152, Medical Center,
University of Virginia,
Charlottesville, Virginia 22908, USA

1. Marks, D. F. *Nature* 320, 119–124 (1986).
2. Koestler, A. *The Roots of Coincidence* (Hutchinson, London, 1972).
3. Rhine, L. E. *J. Parapsychol.* 15, 164–191 (1951).
4. Gurney, E., Myers, F. W. H., & Podmore, F. *Phantasms of the Living* (Trübner, London, 1986).
5. Sidgwick, H. & Committee *Proc. Soc. psychical Res.* 10, 25–422 (1894).
6. Sidgwick, E. M. *Proc. Soc. psychical Res.* 7, 30–99 (1981).
7. Boswell, J. *The Life of Samuel Johnson* 246 (Modern Library, New York, 1931).
8. Bergson, H. *Proc. Soc. psychical Res.* 26, 462–479 (1913).

Trouble on the farm

SIR—"Cheap and nasty trees" (leading article *Nature* 322, 195; 1986) are the backbone of Scandinavian forestry. To suggest that they are not good enough for Britain is a striking example of how widespread is that British malaise responsible for so many of our current problems.

E. NICHOLL

Llangunville, Llanrothal,
Monmouth, NP5 3QL, UK

What is the scientific literature?

Professional people have won a poor reputation for their skill at communicating with each other. The complaint may unfortunately be justified.

By what test are the scientific journals counted as literature? The bare minimum of an answer is that they are collectively referred to in this way by their contributors. Collectively, they also have the quality of permanence; they sit on library shelves for decades on end, and are referred to with reverence by those who contribute to later issues. So much is the bread and butter, or at least the bread and water, of scholarship; there are few complaints at the value of the accumulation of journals as the means by which the record of discovery may be reconstructed. But the same cannot be said of the value of the scientific literature as a means by which information is conveyed between people unable to communicate directly, by dialogue. Why should the general admiration of the way that discoveries roll off the laboratory benches coincide with a general belief that the scientific literature is impenetrable, often to its users?

There are some obvious excuses that do not deserve the dignity of being called explanations. Thus, some hold that the process of discovery inevitably splinters previously coherent fields of study, giving sub-fields an existence and a language of their own and creating barriers over which the practitioners no longer wish to jump. Others say that the content of science steadily becomes conceptually more difficult than can be handled by language as it is, which is a way of saying that the problem is not the fault of professional scientists but of other unspecified professionals. Still others protest that there would be no difficulty if only journals and their editors were more generous with the space they allow to the practitioners who are the creators of this growing volume of literature. All three defences are implausible.

That seems to be a legitimate inference from the opinions of regular readers of the journals, even journals such as this which aim to cater for readers with a variety of interests, both in particular but different fields of research and in more general matters. It is sobering how often those who kindly answer questionnaires about this weekly partial diet of science reply with the opinion that it must all be worthy, but that there are only a few parts of it that they could hope, and even wish, to understand. That there are conceptual difficulties that many people do not have the time to surmount is forgivable. It is even proper that people working in some chosen field should choose to stay there, at least

for the time being. Why should everybody be required to be a generalist? But the sheer clumsiness of the scientific literature is a needless impediment not merely to wider understanding but even to the understanding of those at whom a specialized paper may be directed.

Although the common faults are not easily categorized (and people are forever inventing more), the general drift of error is generally understood even by the transgressors. There is, for example, a disinclination to believe that rules of grammar are particularly relevant to the understanding of scientific prose, whatever their significance in other connections. To judge from the literature as a whole, scientists care less for the accuracy with which the number of a pronoun agrees with that of its noun than they do for the numerical data they report, which is strange to say the least. It is also strange that people who fuss endlessly about the quality of reproduction of a diagram can so fecklessly scatter a word such as "only" in a sentence, forgetting that it belongs only where it most properly qualifies its antecedent.

To judge from much of what is published, most contributors must hope that inventiveness will see them (and others) through. Stringing long lists of words together was the stratagem of the 1970s (as in "language-inventing-capability" or "DNA-binding-protein-sequence-homology"). The stratagem serves no useful purpose because, until the string of words is so familiar within a sub-discipline that it can be represented simply by initials, it remains for the reader to unpack the concatenation each time he comes across it. To how many is RFLP for "restriction-fragment-length-polymorphism" (or is it FLRP for "fragment-length-restriction-polymorphism"?) a concept to hold in one's head as clearly as that of, say, "rabbit"? Even so some journals have now gone so far as to grade the length of the hyphens by which the words are separated in such a string, as if it were always possible to tell which pairs of consecutive words are the more closely related.

Meanwhile, new forms of inventiveness have already begun to appear. Ribose is a sugar which, when combined appropriately with a purine or pyrimidine phosphate, will make a nucleotide; everybody knows that. So why not coin the verb "to ribosylate"? There can be no objection. But what is the object of this ribosylation? ATP, or GTP, or something quite differ-

ent? No problem. The statement suggests the solution. Why not "ATP-ribosylate", in whatever tense may be appropriate? The difficulty that unfamiliarity may often impede the understanding of even those who know what is intended will surely melt away in time, as the whole world takes the trouble to learn this special language?

The difficulty of the literature derives in part from the neglect of many common rules and the accompanying invention of others which are thought preferable, perhaps because they are better suited to the circumstances. But there are faults of a more structural character. The conventions of the research business, for example, require that authors' delight in what they have accomplished should be restricted to the use of conventional phrases, among which "for the first time" is unaccountably one. The result is that an author is prevented from giving an account of why his or her paper is important (which would help towards its understanding) but is allowed baldly to state that an observation has been made for "the first time in a closed vessel under water". Similarly, the convention that a person describing the results of a series of biochemical experiments should give such a full account of the details that a novice starting from scratch could repeat the work unaided is a pleasing concession to the need that work should be reproducible, but no substitute for an intelligent variation on some phrase such as "we followed so-and-so's method with this difference". Yet nobody seems to care that readers must each time unpack each account of an experiment in the literature before they can properly understand its novelty.

Thus the underlying fault with the scientific literature is its intricacy, its requirement of readers that they should not merely read and understand, but read, pull to pieces and then put together again in ways they understand. Given the requirement, the process is plainly more prone to ambiguity than the ordinary person would think credible. And there is nothing in the demands of specialization, or in those made by editors on authors, to suggest that the characteristics of the scientific literature that give it a bad name are externally imposed. This, it seems, is what people want. But is not much of the rest of scholarship much the same? Of course.

John Maddox

this step leads to the activation of the *virG* protein which positively regulates *vir* transcription. Interestingly, *virG* expression is both constitutive and plant-inducible and produces two distinct differentially regulated transcripts: the constitutive transcript is shorter than the induced transcript and lacks sequences required for efficient translation. This feature is reminiscent of the P_{RE} - and P_{RM} -driven transcripts of the CI gene of bacteriophage lambda¹¹.

DNA sequence analysis also suggest that the *virG* locus encodes a 267-amino-acid protein with significant similarity to the *ompR* gene product of *Escherichia coli*¹². The *ompR* gene is needed for activation of expression of the outer membrane porins, encoded by *ompC* and *ompF*, and required for response to osmotic pressure. Quite how the *virG* protein activates transcription of the inducible *vir* loci is not yet clear. It could bind to operator sequences to allow RNA polymerase holoenzyme to initiate transcription. But *virG* appears to be expressed at a significant level, unusual for an activator, and maximal induction of the *vir* loci takes 12–18 hours. Alternatively the *virG* protein could replace the normal vegetative sigma factor in the holoenzyme complex specifically to transcribe the *vir* loci. Those loci, which are exclusively induced by AS, lack the –35 bp consensus sequence in the promoter yet have other hexanucleotide DNA sequences in common around this region, which might function as *cis*-acting regulatory elements for *vir* induction¹³. Further mysteries remain with regard to *vir* regulation; although *virG* is absolutely required for expression of the other inducible loci, AS alone is partially able to induce *virG* transcription even when the *virG* protein is non-functional¹⁴.

What is the mechanism of T-DNA transfer from agrobacteria to plant cells that results from *vir* activation? Initially it was thought that before transfer the T-DNA excises from the Ti-plasmid to form a double-stranded circular DNA molecule joined precisely within one border sequence repeat (by recombination between the two borders)¹⁴. However, this genetic evidence was circumstantial. In the report in this issue¹, Stachel *et al.* provide direct biochemical evidence that AS specifically induces the production of a linear, single-stranded, T-DNA molecule designated the T-strand; such strands are detected by Southern blot analysis of total DNA prepared from induced agrobacteria and are sensitive to different single-strand specific nucleases. These molecules correspond to the bottom strand of the Ti plasmid such that its 5' and 3' ends map to RB and LB, respectively, matching the genetic polarity of the border repeats. T-strand formation is probably associated with DNA synthesis as AS induction also yields molecular changes to

T-DNA sequences in the Ti plasmid — a strand-specific nick associated with the 25-bp border repeats, and Ti-plasmids whose internal T-regions show partial single-stranded and/or triple-stranded character (presumably replicative intermediates). The border nicks probably direct site-specific recombination to yield the double-stranded T-DNA circle molecule observed at low frequency after induction¹⁴.

This evidence suggests the following analogy to bacterial conjugation: after induction of the *vir* regulon by AS, a nick is made within the RB. New data (S.E. Stachel, personal communication) indicate that the function responsible for this nick (and T-strand formation) is encoded by the upstream cistron of *virD*. The *virD* gene is also necessary for the generation of T-DNA circle molecules¹⁵. Rolling-circle DNA replication is then initiated, and in conjunction with a helicase activity, extrudes the T-strand in the 5' to 3' direction. The LB nick signals termination of transfer and, when deleted, results in termination further along. The polar function of RB within the large Ti-plasmid indicates that in the wrong orientation transfer initiates in the other direction and is unlikely to traverse the entire Ti-plasmid to reach the T-DNA (analogous to an interrupted mating); in smaller plasmids the RB is fully functional in both orientations¹⁶. Can transfer initiate from the LB too? Probably, but recent evidence from octopine Ti plasmids indicates that RB is

used more efficiently¹⁷.

Many questions remain to be answered. How is the T-DNA transported from *Agrobacterium* to the plant nucleus? Where is the complementary strand to the T-strand made and what primes its synthesis? How does T-DNA integrate into the nuclear genome? Clearly, further exciting discoveries lie ahead in the understanding of the *Agrobacterium*–plant-cell interaction. □

1. Stachel, E.E., Timmerman, B. & Zambryski, P. *Nature* **322**, 706 (1986).
2. Gheysen, G., Dhase, P., Van Montagu, M. & Schell, J. in *Genetic Flux in Plants* (eds Hohn, B. & Dennis, E.S.) 11 (Springer, Wein, 1985).
3. Douglas, C.J., Staneloni, R.J., Rubin, R.A. & Nester, E.W. *J. Bact.* **161**, 850 (1985).
4. Wang, K., Herrera-Estrella, L., Van Montagu, M. & Zambryski, P.C. *Cell* **38**, 455 (1984).
5. Willetts, N. & Wilkins, B. *Microbiol. Rev.* **48**, 24 (1984).
6. Lichtenstein, C. & Draper, J. in *DNA Cloning* Vol. II (ed. Glover, D.M.) Ch. 4 (IRL, Oxford, 1985).
7. Stachel, S.E., An, G., Flores, C. & Nester, E.W. *EMBO J.* **4**, 891 (1985).
8. Stachel, S.E. & Nester, E.W. *EMBO J.* **5**, 1145 (1986).
9. Stachel, S.E., Nester, E.W. & Zambryski, P.C. *Proc. natn. Acad. Sci. U.S.A.* **83**, 379 (1986).
10. Stachel, S.E., Messens, E., Van Montagu, M. & Zambryski, P.C. *Nature* **318**, 624 (1986).
11. Stachel, S.E. & Zambryski, P.C. *Cell* **46**, 325 (1986).
12. Ebert, P., Stachel, S.E., Winans, S.C. & Nester, E.W. (submitted).
13. Das, A. *et al.* *Nucleic Acids Res.* **14**, 1335 (1986).
14. Koulikova-Nicola, Z., Shillito, R.D. & Hohn, B. *Nature* **313**, 191 (1985).
15. Alt-Moerbe, J., Rak, B. & Schröder, J. *EMBO J.* **5**, 1129 (1986).
16. Rubin, R.A. *Molec. gen. Genet.* **202**, 312, (1986).
17. Peralta, E.G., Hellmiss, R. & Ream, W. *EMBO J.* **5**, 1137 (1986).

Conrad Lichtenstein is in the Centre for Biotechnology, Imperial College of Science and Technology, London SW7 2AZ, UK.

Palaeoclimate

Permafrost and ice-wedge growth

from Peter Worsley

PERMAFROST, any natural material that contains a thermal state of 0°C or less for two or more years, underlies one fifth of the terrestrial surface of the Earth. Its formation is a direct function of climate; its boundaries are determined by a mean annual air temperature of 0 to –1°C. There are few precise data on permafrost establishment, so the study reported in this issue (Peyette, S. *et al.* *Nature* **322**, 724; 1986), demonstrating that the removal of a forest cover, which normally promotes snow accumulation and hence protection from deep freezing, appears to be the key factor causing permafrost growth, is a significant confirmation of the sensitive environmental thresholds governing the onset of permafrost.

Permafrost is generally divided into three types — continuous, discontinuous or alpine — depending on its relative extent in the subsurface. Alpine permafrost is restricted to high mountains and plateaux outside the main polar-centric belt where altitudinal cooling replicates the severe climates associated with the arctic and

subarctic. In the Northern Hemisphere the continuous/discontinuous permafrost areas form two concentric zones with boundaries frequently transverse to latitude because of permafrost sensibility to land-mass configuration. The transition between continuous and discontinuous zones corresponds roughly with a mean annual air temperature of –6 to –8°C.

Cracking of permafrost arises as a consequence of rapid cooling during winter in conjunction with ground temperatures in the –15 to –20°C range. Once initiated, the repeated annual cooling induces cracking of the frozen ground at the same location; successive events over several years, together with attendant infilling of the cracks with blown snow and hoar frost, leads to the slow growth of ice wedges, linear ground-ice bodies with V-shaped cross-profiles. Active growth of ice wedges has until recently been considered to be an exclusive property of continuous permafrost and, although ice wedges in discontinuous permafrost were known, they were regarded as being dormant



Medium-sized ice wedges on an actively eroding coast in Banks Island, Canada (72°N, 130°W) in continuous permafrost (mean annual air temperature is about -15°C).

because of a climatic amelioration. Hence the wedges were thought to be relics from a former more severe, climatic regime that could sustain continuous permafrost in the subsurface. The reappraisal of this view by Peyette *et al.*, leading to the conclusion that active ice wedges do occur within discontinuous permafrost, was prompted by the study of T.D. Hamilton *et al.* (*Arctic Alp. Res.* **15**, 157; 1983) who, as a result of trench excavation during design of the trans-Alaskan pipeline, demonstrated ice-wedge growth in a mean annual temperature close to -3.5°C. The vegetation cover in the study of Hamilton *et al.* was similar to that which prevails in the interior subarctic boreal forest of Quebec.

Peyette *et al.* studied a site near Clearwater Lake in subarctic Quebec which has a mean annual air temperature of -5°C and extensive discontinuous permafrost. The association of ice wedges and peat allows radiocarbon isotope-dated, palaeoecological reconstructions of vegetation history the latter being the key to their demonstration that ice-wedge initiation was a consequence of the removal of tree cover. The relationships between ice-wedge width and distance from the contemporary forest border uncovered by Peyette *et al.* are an elegant proof of the diachronous nature of ice-wedge initiation. Without the progressive retreat of the forest border no growth would have occurred, which emphasizes the need for caution in the deduction of climatic significance from such structures. But the host materials are highly organic and a low preservation potential is likely if the permafrost should degrade in the future.

Wedge-shaped sedimentary structures

in the geological record are usually considered to be casts arising from the melt of ice wedges. Such melting can be caused by regional climatic change or local factors such as thermal erosion of the ice by melt water during the summer. They are very valuable indicators of the former presence of permafrost and thus important aids to palaeoclimatic reconstruction. Climatic parameters derived from ice-wedge casts are usually based on the assumption that ice-wedge growth signifies continuous permafrost and therefore mean annual temperature of at least -6°C. No active

ice wedges have yet been found that originate in non-organic sediments within a zone of discontinuous permafrost. But in the light of the findings of Peyette *et al.* and Hamilton *et al.*, a cautious re-assessment of the climatic values that control ice-wedge initiation and subsequent growth is necessary: factors such as rates and degree of cooling of the upper permafrost must now be considered. □

Peter Worsley is in the Department of Geography, University of Nottingham, Nottingham NG7 2RD, UK.

Immune system

Complement-like cytotoxicity?

from Kenneth B.M. Reid

BOTH cytotoxic T cells and natural killer cells have cytoplasmic granules containing the protein perforin, which appears to have an analogous function to that of serum complement component C9 in the lysis of target cells^{1,2}. Activation and regulation of the complement system is quite well understood but little is known about the system that controls perforin activity. Efficient and controlled lysis of target cells by the complement system involves six serine proteases³. It might be expected that natural killer and cytotoxic T cells contain a complex complement-like regulatory system and recent studies at the protein and complementary DNA levels, one of which appears on page 740 of this issue, indicate that this is the case as there appear to be at least two serine proteases involved in cytotoxic T-cell activity.

Pasternack *et al.* find that a serine protease is involved in the antigen-specific lysis of target cells by mouse cytotoxic T cells (and certain helper T cells). The principal protease identified is composed of two, probably identical, disulphide-linked subunits of approximate relative molecular mass (M_r) 32,000 (32K). Inhibition of this protease by diisopropyl-phosphorofluoridate and its limited trypsin-like activity indicate that it belongs to the serine protease family. The enzyme appears to be associated with the membranes of the perforin-containing cytoplasmic granules and its release is specifically triggered by the interaction of cytotoxic T cells with target cells.

Studies at the cDNA level by the isolation of clones specific for cytotoxic T cells predict protein sequences for three molecules having the characteristic features of the classical serine proteases, that is, a chain of $M_r \approx 25K$ containing the highly conserved residues, such as those at the active catalytic site, His, Asp and Ser. Brunet *et al.* reported the isolation by subtractive hybridization of a mouse cDNA clone (CTLA-1) specific for cytotoxic T

cells but not helper T cells or B cells. The sequence predicts a leader peptide of up to 20 amino acids followed by a sequence of 227 residues resembling a serine protease beginning with the characteristic Ile-Ile-Gly-Gly amino-terminal sequence.

An identical cDNA sequence has been identified by Lobe *et al.* who first found two genes encoding serine proteases that appeared to be specifically expressed in activated cytotoxic T cells, just before the cytotoxic reaction. But on screening cDNA libraries using the genomic probes, coding sequences for only one of the proposed proteases (CCP1) could be isolated. In another study the cDNA encoding a serine protease-like sequence different to that seen for CTLA-1/CCP1 clones, was isolated from a mouse cytotoxic T-cell cDNA library⁷ using an RNA hybridization competition procedure. The sequence of this clone predicted an amino-terminal sequence of 14 amino acids (which seems likely to be part of an activation peptide rather than a leader sequence) followed by 232 amino acids of serine protease-like sequence. The molecule identified (HF) was found in all cytotoxic T-cell clones, nude mice spleen cells and rat natural killer cells examined but not in control tissues and cell lines.

The derived amino-acid sequences for the two serine protease-like molecules CTLA-1/CCP1 (refs 5,6) and HF (ref. 7), show 40 and 35 per cent homology with rat mast cell protease II, an intracellular serine protease found in the granules of atypical mast cells⁸. Human factor D (refs 9,10), the first enzyme involved in the activation of the alternative pathway of complement, shows 36 per cent homology with the rat mast cell protease⁹ and appears to be the most closely related of the complement serine proteases to the mouse cytotoxic T-cell-specific proteases. The presence of an Asp residue in the position determining the substrate specificity of the HF protease indicates that, like

factor D and the other complement enzymes it would have a preference for lysyl or arginyl bonds. This is consistent with the observation of Pasternack *et al.*¹ that the protease they describe has a limited trypsin-like activity.

The presence of Ala in the equivalent position in the other cytotoxic T-cell-specific protease, CTLA-1/CCP1, makes this likely to be more similar to the rat mast cell protease II in its activity. The CTLA1/CCP1 protease further resembles the mast cell protease in that it has Phe, rather than the highly conserved Cys seen in all other serine proteases, at the position four residues on the amino-terminal side of the serine at the active site, and this could affect the substrate binding pocket. Also, the presence of an uneven number of Cys residues in the derived amino-acid sequence for the CTLA-1 and CCP1 sequences (and perhaps HF) implies that the Cys at position 74 has a free sulphhydryl group, or is linked by a disulphide bond to another chain, as it has no homologous counterpart in all other known serine protease sequences. It is possible that the disulphide bond in the dimer observed by Pasternack *et al.*¹ involves this Cys residue. This would certainly be a novel observation and could be a feature of cytotoxic T-cell-specific serine proteases.

Thus it is possible that the CTLA-1/CCP1 and HF clones encode two of several distinct serine proteases in a lytic pathway common to natural killer and cytotoxic T cells. Cytotoxic T cells perhaps follow a classical antigen-mediated pathway whereas natural killer cells may use an alternative pathway not involving the T-cell receptor, with both pathways culminating in a common lytic process. If there is close homology between the lytic processes of these two cell types and that of the serum complement system then proteins homologous with the complement components C3, C4 and C5 may be found acting as substrates for these recently identified serine proteases, thus allowing generation of complex proteases similar to C4b2a3b and C3bBb3b that are involved in the generation of C5b in the serum complement system. □

1. Podack, E.R., Young, I.D. & Cohen, Z.A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8629 (1986).
2. Henkart, P.A., Millard, P.J., Reynolds, C.W. & Henkart, M.P.J. *exp. Med.* **160**, 75 (1984).
3. Reid, K.B.M. & Porter, R.R. *A. Rev. Biochem.* **50**, 433 (1981).
4. Pasternack, M.S., Verret, C.R., Liu, M.A. & Eisen, H.N. *Nature* **322**, 740 (1986).
5. Brunet, J. F. *et al.* *Nature* **322**, 268 (1986).
6. Lobe, C.G., Finlay, B.B., Paranchych, W., Paetkau, V.H. & Blackley, R.C. *Science* **232**, 858 (1986).
7. Gershenfeld, H.K. & Weissman, I.L. *Science* **232**, 854 (1986).
8. Woodbury, R.G., Katunuma, N., Kobayashi, K., Titani, K. & Neurath, H. *Biochemistry* **17**, 811 (1978).
9. Johnson, D.M.A., Gagnon, J. & Reid, K.B.M. *FEBS Lett.* **166**, 347 (1984).
10. Niemann, M.A., Bhowan, A.S., Bennett, J.C. & Volanakis, J.F. *Biochemistry* **23**, 2482 (1984).

Kenneth B.M. Reid is in the MRC Immunochimistry Unit, University of Oxford, Oxford OX1 3QU, UK.

Membrane biophysics

Surface conduction of protons

from M.J. Selwyn

IN AN elegant series of experiments on phospholipid monolayers, Prats, Teissie and Tocanne, on page 756 of this issue¹, report that protons move much more rapidly in a thin layer (less than 0.5 nm thick) at the interface between the lipid and aqueous phases than through the thickness of the aqueous phase (a few millimetres). This model system provides experimental evidence for a reaction-facilitating function of membranes and is particularly interesting in relation to chemiosmotic coupling by protons. The result also has implications for the study of reactions at non-biological interfaces, such as catalysis by immobilized enzymes.

The Langmuir trough technique, used since 1917 for measuring the surface pressure and the cross-sectional area of the constituent molecules of monolayers

is that they can measure this parameter for any phospholipid without the perturbations caused by an added probe group.

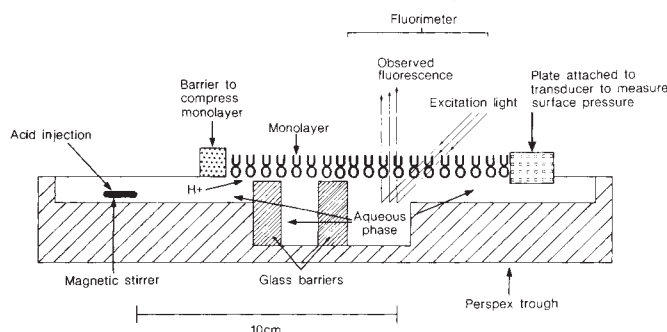
They find that reporter-group signals change more rapidly than the pH of the aqueous phase and that there is both a pH gradient along the surface of the membrane and a steep pH gradient from the interface through a thin layer (measured by the linked fluorescein, estimated to be 1.0 ± 0.5 nm into the aqueous phase) to the bulk aqueous phase. Thus, fast diffusion along the interface and slower diffusion into the bulk aqueous phase compete for interfacial protons. Fast diffusion disappears if the lipids are not sufficiently tightly packed in the monolayer, which suggests some structural requirement. But increasing the ionic strength with NaCl or thiocyanate does not alter

the rates of diffusion², which indicates that 'structured' water is not involved. In contrast, increasing the concentration of K-phosphate in the buffer increases trapping of the protons by the aqueous phase so much that conduction along the interface cannot be detected³.

Although it is now generally agreed that electron transfer and ATP synthesis in oxidative and photosynthetic phosphorylation are coupled by the flow of protons, some quantitative data do not fit

with P. Mitchell's original concept that the protons equilibrate with the bulk aqueous phases and others exclude the alternative hypothesis of R.J.P. Williams in which the protons are confined to the membrane. Since then there have been a plethora^{4,7} of hypotheses suggesting semi-localized proton coupling in which the resistance to transfer of protons to the bulk phase is neither infinite (as in Williams hypothesis) nor zero (as in Mitchell's).

The main points of contention are whether there are structural barriers to proton flow to and from the bulk aqueous phases; whether there are specific channels between the proton-linked enzymes; and the extent of localization of the protons, ranging from the whole surface of each vesicle membrane down to units comprised of single sets of the proton-



Diagrammatic longitudinal section of a Langmuir trough modified for injection of acid and measurement of fluorescence changes at and near the surface (from refs 1-3). Fluorescence is detected both in the monolayer and in the aqueous phase just below it. Thus, depending on whether the fluorescein pH indicator is attached to the lipids in the monolayer or free in solution the fluorimeter responds to the pH at the interface or to the pH of the bulk aqueous phase, respectively. Controls are provided by pH measurements at the same position in the bulk aqueous phase without the monolayer or the glass barriers. The shorter time-constant for pH change when the monolayer is present demonstrates rapid conduction of protons at the interface.

spread on a water surface, is adapted by Prats *et al.* so that they can inject acid at one end of the trough and, further along, measure changes in the aqueous pH and signals from chemical groups attached to the monolayer (reporter groups) which have a readily measurable property that responds to changes in the local environment (see figure). In previous experiments^{2,3} they found the pH close to the lipid-water interface by measuring changes in the fluorescence of fluorescein that is covalently linked to phosphatidylethanolamine. In their new work they use americium electrodes to measure changes in the surface potential of the lipid monolayer and thus the state of ionization of the lipid headgroups. This ionization depends on the pH precisely at the interface and one advantage for Prats *et al.*

linked enzymes. Calculations suggest⁶ that even in the aqueous phase the movement of protons should be so rapid that without special barriers the protons have rapid access to the whole surface of the membrane and to the bulk aqueous phases. But the results of Prats *et al.* indicate that proton movement away from a phospholipid membrane is not rapid, even in the absence of special barriers, and so steady-state fluxes of protons between sources and sinks on the membrane surfaces and to the internal and external aqueous phases should be considered.

On this basis the bulk-phase pH and membrane potential are not equal to the local values either for the reduction-oxidation enzymes or for the ATPase, but there is no definite distinction between localized and delocalized coupling. An analogy can be made with a spring or fountain where a collector placed close to the source can collect water at a higher potential energy than that in the surrounding pool. Unless diffusion of protons across the surface of the membrane is infinitely fast then there will be a progressive tendency for proton-linked enzymes to be more tightly coupled and more isolated from the bulk phases the closer they are on the membrane surface. With this type of situation, depending on conditions, the system may tend either towards localized or delocalized behaviour, and indeed Beard and Dilley recently showed such a change by storing thylakoids in high rather than low ionic-strength medium.

Adam and Delbrück⁹ proposed that intracellular membranes enhance reactions rates by constraining diffusion to two dimensions on the surface of the membrane. This hypothesis has been criticized, partly on the grounds that a great reduction in diffusion rate at the surface offsets much of the advantage. But the rapid movement of protons at the membrane surface may affect the rate of many reactions and maintain uniformity of cytoplasmic pH. Although monolayer experiments are a very simple model, they should also stimulate investigation of other ions and non-biological surfaces. □

1. Prats, M., Teissie, J. & Tocanne, J. F. *Nature* **322**, 756 (1986).
2. Teissie, J., Prats, M., Soucaille, P. & Tocanne, J. F. *Proc. natn. Acad. Sci. U.S.A.* **82**, 3217 (1985).
3. Prats, M., Tocanne, J.-F. & Teissie, J. *Eur. J. Biochem.* **149**, 663 (1985).
4. Mitchell, P. *Biol. Rev.* **41**, 445 (1966).
5. Williams, R. J. P. *Biochim. biophys. Acta* **505**, 1 (1978).
6. Westerhoff, H. V., Melandri, B. A., Venturoli, G., Azzone, G. F. & Kell, D. B. *Biochim. biophys. Acta* **768**, 257 (1984).
7. Ferguson, S. J. *Biochim. biophys. Acta* **811**, 47 (1985).
8. Beard, W. A. & Dilley, R. A. *FEBS Lett.* **201**, 57 (1986).
9. Adam, G. & Delbrück, M. in *Structural Chemistry and Molecular Biology* (eds Rich, A. & Davidson, N.) 198 (Freeman, New York, 1968).
10. McCloskey, M. A. & Poo, M. J. *Cell Biol.* **102**, 88 (1986).

M. J. Selwyn is in the School of Biological Sciences at the University of East Anglia, Norwich NR4 7TJ, UK.

Astronomy

Neutron stars in Nanjing

from Virginia Trimble

NEUTRON stars form in supernova explosions, manifest themselves as radio pulsars or X-ray binaries, and fade away as their rotation slows, their magnetic fields decay away and their energy supplies are used up. This basic framework has been accepted for nearly 15 years, so that one might wonder whether there was anything left for a conference on the origin and evolution of neutron stars to discuss.

But several fundamental questions (and a host of minor ones) remain. Are there other ways of making neutron stars? And, perhaps most important, when and how do they acquire, then lose, the strong dipole magnetic fields and rapid rotation that permit pulsar emission and provide signatures of neutron-star presence in X-ray sources? Unsurprisingly some possible answers are emerging from studies of two very rare classes of pulsars — those with rotation periods less than 10 ms and those with compact (white dwarf or second neutron star) binary companions. The classes overlap in that two of the three known millisecond pulsars are among the seven known binary ones. Two papers by Joseph H. Taylor and colleagues on pages 712 and 714 of this issue report the most recent additions to each group. More can be expected, for they probably make up about 10 per cent of all galactic pulsars, according to Taylor (Princeton University) and Ramesh Narayan (University of Arizona), and are currently under-represented in catalogues only because their intrinsic faintness and rapid pulsation makes detection difficult.

What are the binary and millisecond pulsars telling us? First, they provide the firmest evidence so far for neutron-star formation not from the collapse of the core of a massive star, but rather from collapse of a roughly solar mass white dwarf driven above the Chandrasekhar (maximum stable) mass by accretion from a companion.

This mechanism, originally invoked to account for young-looking neutron stars with 10^{12} Gauss magnetic fields among old, low-mass X-ray binaries, seems to be the only way to explain the combination of young neutron star (10^{11} Gauss surface field) and old system (wide, circular orbit) that characterizes the binary pulsar 0820+02 (ref. 4). A recent detection of the companion⁵ strengthens the explanation. The other star is a hot white dwarf, which must have ceased transferring material to its companion about 10^7 years ago, the elapsed time being just right for the white

dwarf to cool to its present temperature and for the neutron star to have reached its present field and rotation period (0.865 s).

Initially strong fields and rapid rotation are expected simply from conservation of magnetic flux and angular momentum as stellar cores collapse to neutron stars. Statistics of pulsar periods, period changes, and associations with supernova remnants tell us, however, that only about 20 per cent are born with the largest possible values of both field and angular velocity. Stanford E. Woosley (University of California, Santa Cruz) suggested that some neutron stars might be born as slow rotators from progenitor stars larger than 10 solar masses, because core convection during carbon and silicon burning could transport angular momentum outward. Lower masses without such convection would give rise to rapid rotators.

The alternative of low initial surface field must also sometimes occur, for instance in LMC X-4, whose massive companion requires the neutron star to be young. Given the 69-ms rotation period, the accretion that fuels observed X-ray emission can only occur with a surface field less than 3×10^{10} Gauss, according to Nicholas E. White (Darmstadt). If, when



100 years ago

The August Perseids

THE shower of Perseids has been a fairly conspicuous one this year notwithstanding the somewhat unfavourable circumstances attending the display. On the nights of August 9, 10, and 11 the nearly-full moon was visible during the greater part of the time available for observation, and robbed the phenomenon of its chief prominence during the evening hours. Those, however, who continued to watch the heavens until after the moon set on the early morning of the 11th must have been rewarded by a tolerably rich exhibition of meteors. The number observable by one person fell little short of 100 per hour, and this rate compared with similar observations in past years proves the late display to have fully maintained its decided character. Numerically this shower of Perseids cannot be placed in the same category as the brilliant meteoric storms of November 13, 1866, and November 27, 1872 and 1885, but it must be remembered that the August shower is one which returns *annually*, and apparently without much variation in its leading features.

From *Nature* **34**, 372; 19 August 1886.

*IAU Symposium No. 125 was held in Nanjing, China, 26–30 May 1986.

or how such initially low fields grow to typical values of 10^{12} Gauss is not certain, though mechanisms have been suggested⁶.

In any case, pulsars then live on their rotational (and, perhaps, magnetic) energy until dropping below detectability at periods of a few seconds and affective fields of 10^{11} Gauss, preventing the study of later phases. One catch is that only the dipole part of the field perpendicular to the rotation axis is measured, so that one could not rule out the field aligning or evolving to higher multipoles rather than decaying in the 10 – 10^7 years over which pulsars fade away. The binary pulsars have opened a new window for observation of senile neutron stars. As the companion evolves and expands, gas transferred to the neutron star spins it back up to an equilibrium rotation period, depending somewhat on orbit parameters and field strength, but typically a few milliseconds. Later, when the companion itself has become compact or unbound and ceased to interfere, we see a recycled pulsar with rapid rotation and low field.

One might expect the magnetic field strength to continue to drop. Apparently it does not⁷. All three millisecond pulsars have fields very close to 5×10^8 Gauss, according to Taylor. This includes the binary 0655+64, whose white dwarf companion has been cooling for several billion years since it stopped spinning up the neutron star. Corroborative evidence for stalling of field decay at 10^8 – 10^9 Gauss comes from models for quasi-periodic oscillations in old (low mass) X-ray binaries⁸, although so many different kinds of correlations of the frequencies of these oscillations with X-ray intensity and colour temperature have now been seen that the connection with any one model has become very tenuous according to Guenther Hasinger (Max Plank Institute, Munich).

Apparently, then, we can add to the basic framework mentioned above an alternative mechanism (accretion-driven white dwarf collapse) and alternative evolutionary histories for rotation and surface magnetic field more complex than simple monotonic decay with time.

Other neat, new things that were reported at Nanjing included a report from Andrew Lyne (Jodrell Bank) of the largest glitch ever, $\Delta P/P = 4 \times 10^{-6}$, which occurred in January or February 1986 in PSR 0355+534 ($P = 0.156$ s). The recovery has also been unprecedentedly rapid ($P = -52 \times 10^{-24}$ s⁻¹). One might or might not want to tie this up with the suggestion by Charles Alcock (MIT) that post-glitch behaviour is a possible test to distinguish true neutron stars from interlopers made of strange quark matter. Claire Flanagan (SAAO) described a new glitch in the Vela pulsar almost as large ($\Delta P/P = 10^{-6}$).

Jichao Huang (Nanjing University)

advocated dividing pulsars into two roughly equally numerous groups that emit at polar caps and at their speed-of-light cylinders respectively; and Tan Lu (Nanjing University) had statistical evidence that the fading of pulsar radio luminosities may not be monotonic with time.

Patrizia Caraveo and Giovanni Bignami (Milan) have added another late-1985 EXOSAT point to the curve of period as a function of time for the 60 s pulsation of Geminga. The object is still slowing down and now shows two harmonics as well as the fundamental. Caraveo and Bignami have not, however, been able to get additional optical images to test the possibility of large proper motion. Zhenru Wang (Nanjing University) suggested several new possible identifications between Chinese guest stars and young neutron stars and supernova remnants (including Geminga!).

W. David Arnett (Chicago) has found a new instability following core collapse in massive stars that may advect enough energy and momentum outward to cause

envelope ejection and a supernova explosion. The most recent round of neutron-star cooling calculations requires pion condensation, strange quark matter or some other non-standard physics to prevent the Vela pulsar from emitting more thermal X rays than the observed upper limit, according to Ken'ichi Nomoto (Tokyo) and Sachiko Tsuruta (Montana State University). These are just a few of the facts presented at Nanjing, which it is hoped will form the raw materials for a new synthesis. □

1. Dewey, R.J. *et al.* *Nature* **322**, 712 (1986).
2. Segelstein, D.J. *et al.* *Nature* **322**, 714 (1986).
3. Whelan, J.A.J. & Iben, I. *Astrophys. J.* **186**, 1007 (1973).
4. van den Heuvel, E.P.J. & Taam, R.E. *Nature* **309**, 235 (1984).
5. Kulkarni, S. *Astrophys. J. Lett.* **306**, L85 (1986).
6. Blandford, R.D., Applegate, J.H. & Hernquist, L. *Mon. Not. R. astr. Soc.* **204**, 1025 (1983).
7. Bhattacharia, D. & Srinivasam, G. *Curr. Sci.* **55**, 327 (1986).
8. Alpar, M.A. & Shaham, J. *Nature* **316**, 239 (1985).
9. Romani, R. & Trimble, V. *Nature* **318**, 230 (1985).

Virginia Trimble is at the University of California Irvine California 92717 and is in the Astronomy Program, University of Maryland, Maryland 20742, USA.

Immunology

The ins and outs of antigen processing and presentation

from Ronald N. Germain

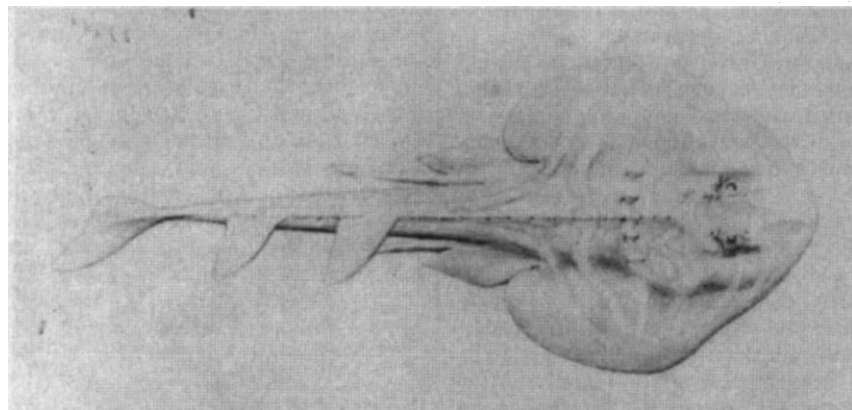
THE striking differences in antigen reactivity shown by T and B lymphocyte subpopulations has been the primary research interest of many members of a whole generation of immunologists. The expectation that elucidation of the structure of the T-cell receptor would, by itself, reveal the how and why of the differences has proved false. Therefore, increasing attention has been paid to understanding the form of antigen actually involved in the T-cell recognition process. The studies of Townsend and his collaborators, culminating in a recent paper¹, provide a major advance in our understanding of the form of antigen seen by T cells and provide an intellectually and physiologically satisfying explanation for several poorly understood immunological phenomena.

Although very similar gene segments and DNA rearrangement events are involved in generating the receptors used by T and B cells, the functional specificity of these two cell types differs markedly. B cells express surface membrane versions of serum immunoglobulin molecules that directly bind soluble antigens of diverse chemical composition. In contrast T cells use a recently characterized heterodimer to co-recognize particular combinations of protein antigen and class I or class II major histocompatibility complex (MHC)-encoded molecules on the surface membrane of other cells.

As early as 1958, Gell and Benacerraf pointed out that antibodies tend to be specific for native protein molecules, whereas cell-mediated (now known to be T-cell-mediated) responses such as delayed-type hypersensitivity are stimulated by determinants preserved after protein denaturation. Pioneering work by Ziegler and Unanue demonstrated that antigen presentation to class II MHC-restricted T cells is an active time-dependent process involving an acidified intracellular compartment whose function is sensitive to drugs such as chloroquine. These results suggested that the antigen seen by T cells is not intact, but rather degraded or denatured intracellularly before exposure on the surface of the antigen-presenting cells. Shimonkevitz *et al.* provided direct evidence in support of this hypothesis by showing that proteins enzymatically degraded *in vitro* could be presented by metabolically inactive cells bearing the proper class II molecules.

Thus, physiological priming of class II restricted T cells to degraded forms of antigen accounts for the presence of secondary T-cell responses to protein antigens denatured *in vitro* and for the success many have had using protein fragments or synthetic peptides as stimuli for T cells originally generated by immunization with intact proteins. Both indirect and direct methods have now revealed that at

Pictures from the *Endeavour* voyage of 1768



Trygonorhina fasciata Müller and Henle, 1841. Drawing by Spöring of a fish caught on 29 April 1770 at Sting Ray Bay (later Botany Bay), Australia. *T. fasciata* was based by Müller and Henle partly on this drawing, but also on another specimen in alcohol at the Paris museum. From *Catalogue of the Natural History Drawings commissioned by Joseph Banks on the Endeavour Voyage 1768–1771 Part 3: Zoology*, recently published by the Bulletin of the British Museum (Natural History). (Historical series vol. 13; London, 1986.) This catalogue lists all the drawings of animals from James Cook's voyage on the *Endeavour* that were made by artists employed by Joseph Banks. Most of the drawings were kept in Banks' home until his death in 1820 and then transferred to the Museum. This catalogue attempts to record comprehensively the surviving animal drawings from the voyage of the *Endeavour*, and examines and discusses the history of the drawings as well as providing notes on the artists, none of whom survived the voyage.

least some of these minimal antigens interact physically with class II molecules in the absence of the T-cell receptor, providing an understandable picture of how simultaneous recognition of the two is achieved.

This view of antigen presentation to class II-restricted T cells diverges markedly from the present view of antigen recognition by class I-restricted T lymphocytes, especially cytotoxic T lymphocytes (CTL). These cells seem specific for integral cell-membrane molecules such as viral envelope glycoproteins, allogeneic MHC molecules, or minor histocompatibility antigens that have been assumed (without any experimental evidence) to be uncharacterized membrane proteins. Immunologists presumed that these intact membrane proteins interacted directly with class I MHC molecules on the cell surface. However, such a marked difference between class I and class II antigen interaction became less and less likely as accumulating data showed striking similarities between the relationship of both class I and class II molecular structures to T-cell recognition¹³, and the use of the same V_{α} and V_{β} gene segments by the receptors of class I- and class II-restricted T cells¹⁴.

Experiments

The experiments of Townsend *et al.*⁵ now reveal that these differences are more imagined than real, and that similar (though not necessarily identical) forms of antigen participate in recognition by all MHC-restricted T cells. The new findings

come from work on the specificity of both human and murine CTL generated by immunization with the influenza virus. Early studies² demonstrated many CTL clones that recognized specificities shared by distinct serotypes of this virus. The viral protein accounting for this cross-reactivity was, unexpectedly, the nucleoprotein and not the membrane-expressed haemagglutinin molecule. Confirmation of these results was obtained using L cells co-transfected with the proper class I gene and the nucleoprotein gene. These cells were recognized appropriately by the cross-reactive CTL clones³.

Given that interaction of CTL receptors with target cells occurs at the outer surface of the cell membrane, how can a molecule that does not appear on the membrane serve as a target antigen? This conundrum was partially solved by the demonstration that transfectants expressing truncated forms of the nucleoprotein molecule could serve as adequate CTL targets⁴. Finally, the experiments described in their most recent paper⁷ show that, just as synthetic peptides could be used to stimulate class II-restricted T cells, target cells exposed *in vitro* to short synthetic peptides corresponding to certain regions of the nucleoprotein molecule could be recognized by nucleoprotein-specific CTL. The active peptides, as would be expected, correspond to regions of nucleoprotein sequence of the immunizing strain of virus that differ from the sequence of nucleoprotein of a strain of virus incapable of sensitizing cells for recognition by these

particular CTL clones. Thus, it appeared unnecessary for the intact nucleoprotein mysteriously to reach the cell surface. Instead, some fragmented or degraded form probably serves as the antigen after travelling to the membrane by a mechanism independent of the normal intracellular sorting process acting on the intact nucleoprotein.

Similarities

Thus, we now have a satisfying similarity in the general nature of antigen recognized by both class I- and class II-restricted T cells. The new data on CTL recognition also provide an explanation for the repeated failure of investigators to generate antibodies to minor histocompatibility antigens. Rather than constituting a family of intrinsic cell-membrane molecules, these antigens are more likely to represent processed fragments of cytoplasmic or nuclear proteins that normally never reach the membrane in an intact state, just as with influenza nucleoprotein. Even if antibodies were raised in response to the very small amount of such fragments that probably are present on the cell surface, they are not likely to be detected in binding or cytotoxicity assays.

The studies on influenza-specific CTL support the emerging view that antigen processing is not a specialized function of a small subset of class II-bearing, bone marrow-derived cells, but rather the reflection of a more fundamental set of intracellular activities occurring in virtually all cells. An early hint that this might be the case came from experiments showing that fibroblasts made to express class II antigens by DNA-mediated gene transfer could also present various protein antigens to T cells, and that this function is susceptible to the same inhibitors that interfere with processing by haematopoietic-presenting cells¹⁵. The demonstration by Townsend *et al.*⁵ that class I-restricted T cells are specific for antigen fragments suggests that most, if not all, cells have some capacity for antigen processing, because class I molecules are widely distributed on various somatic cell types, and CTL recognize virally infected or minor histocompatibility antigen-bearing cells of diverse tissue origin.

Is the same intracellular pathway used for processing those antigens taken up from the external environment of the cell and those synthesized endogenously? Do both classes of MHC molecules interact with or present the same processed antigens? Recently published work by Morrison *et al.*¹⁴ suggests the answer to both of these questions may be 'no'. These investigators find that influenza haemagglutinin synthesized endogenously can not be recognized by class II-restricted, haemagglutinin-specific T-cell clones, but provide an adequate target for class I-restricted cells. Conversely, the class I-restricted

clones ignore class II-bearing cells that had taken up haemagglutinin protein from the culture medium, which are excellent stimulators of class II-restricted lymphocytes. Furthermore, only the class II-dependent antigen presentation is sensitive to the drug chloroquine. Thus, there may be at least two distinct mechanisms for the production of immunogenic protein fragments, and there may be different pathways of intracellular transport for class I and class II molecules that determine the nature of the antigens with which they are recognized.

Model

A speculative scheme of antigen processing and presentation to T cells that takes into account these new data, the demonstration of binding between antigenic peptides and MHC molecules¹², and the recent intriguing report by Cresswell¹⁵ on the presence of class II molecules within the endocytic pathway, is given in the figure.

For exogenous protein antigens, the universally distributed endosomal recycling pathway¹⁶ is considered to have the central role, accounting for the ability of diverse cells to carry out antigen processing. Intact proteins enter the cell by either (fortuitous) binding to molecules engaged in receptor-mediated endocytosis, or as a result of pinocytosis. As the endosomes acidify during their intracellular passage,

this change itself, or the activation of acid-pH-dependent, cathepsin-like proteases leads to the denaturation and/or fragmentation of the original protein molecules. A subset of the resultant peptides, perhaps already associated with class II MHC products, is transported to the cell surface in vesicles bringing molecules such as discharged transferrin receptors back to the membrane. The remaining peptides may go to the lysosomal compartment for complete degradation.

For endogenous proteins, processing may occur in a region of the transitional Golgi specialized for dealing with improperly folded proteins synthesized by the cell. Denatured or fragmented molecules in this compartment may escape sorting into secondary lysosomes, and be transported to the cell membrane along or together with class I molecules.

In both cases, selection of peptides for surface expression presumably depends on structural features that either enhance binding to MHC molecules or other potential carrier proteins cycling to the cell membrane, or that reduce the potential for transport to lysosomes. The α -helical state that is hypothesized to be an important feature of immunogenic segments of a protein molecule may be critical in this regard^{17,18}. Once it is on the cell membrane, the antigen fragment would be available for recognition by the

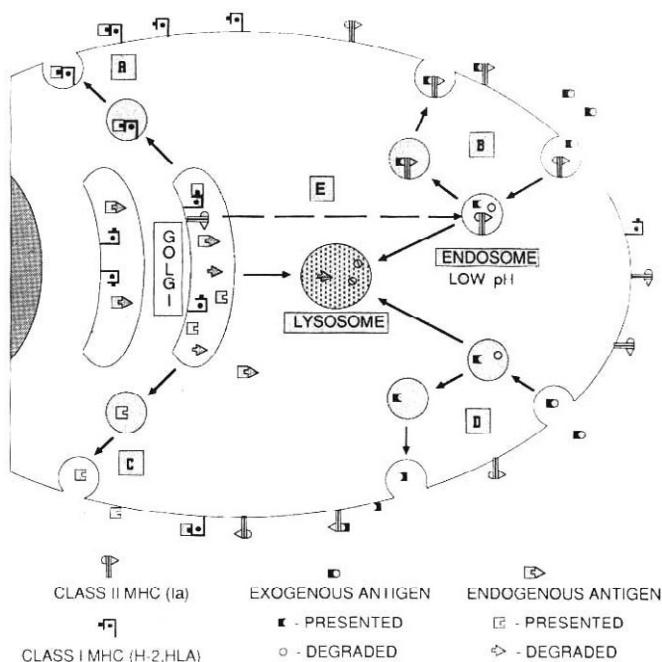
T cell in conjunction with the appropriate MHC molecule.

Why did such complex schemes of antigen processing and presentation evolve? It has been suggested that antigen denaturation or fragmentation is necessary to create molecules with increased lipophilicity, thereby stabilizing interaction with the plasma membrane and increasing the effective local concentration of antigen in the environment of the MHC molecules. In addition, if direct binding of antigen to MHC molecules is essential for T-cell recognition, then proteins that in their native state are much larger than MHC molecules may need to be fragmented to be effective immunogens.

What is particularly apparent in this emerging story is that phenomena considered to be primarily of immunological interest turn out to involve processes of substantial general importance to cell biology. Elucidation of the intracellular events involved in antigen processing will have broad import for our understanding of the traffic patterns followed by proteins within cells. The contributions of Townsend *et al.* remove a major conceptual stumbling block to research in this area and progress should now occur at an accelerating rate. It is also clear that this new picture of antigen recognition by CTL has important implications for the rational design of synthetic vaccines.

Finally, one should not forget the original question — what does the T cell see? With a solution of the crystal structure of HLA close at hand, the capacity to identify the minimal antigenic peptide recognized by a specific T cell together with this class I MHC molecule may provide the means to create a physical picture of the antigen-MHC molecular complex bound by the T-cell receptor. □

The illustration shows a schematic view of the several possible pathways followed by antigen and MHC-encoded molecules in an antigen-presenting cell. Endogenous antigens, such as viral glycoproteins or nucleoprotein, are synthesized in the endoplasmic reticulum and transported to the Golgi where they are shown in an intact and partially degraded form. The peptide representing the immunogen then reaches the cell surface either 1) already associated with class I (H-2, HLA-A,B,C) molecules also exiting from the Golgi in exocytic vesicles (Pathway A, top left), or 2) alone (Pathway C, bottom left), following which the peptide associates with class I molecules already on the cell membrane. For exogenous antigens presented in the context of class II molecules, intact antigen enters the cell by means of endocytic vesicles. As these acidify and partial proteolysis occurs, the immunogenic peptides are associated either 1) with class II molecules recycling from the cell surface in the same endosomes (Pathway B, upper right); 2) with newly synthesized class II molecules from the Golgi (Pathway E, centre); or 3) with class II molecules on the membrane, following transport of the peptide to the cell surface in exocytic vesicles (Pathway D, lower right). In all cases, other peptides are shown here as being transported directly into lysosomes for complete degradation, never reaching the membranes, and hence are unavailable to serve as immunogens.



- Robertson, M. *Nature* 317, 768 (1985).
- Townsend, A.R.M. & McMichael, A.J. *Prog. Allerg.* 36, 10 (1985).
- Townsend, A.R.M., McMichael, A.J., Carter, N.P., Huddlestone, J.A. & Brownlee, G.G. *Cell* 39, 13 (1984).
- Townsend, A.R.M., Gotch, F.M. & Davey, J. *Cell* 42, 457 (1985).
- Townsend, A.R.M. *et al.* *Cell* 44, 959 (1986).
- Kronenberg, M., Siu, G., Hood, L.E. & Shalini, N. A. *Rev. Immun.* 4, 529 (1986).
- Schwartz, R.H. *A. Rev. Immun.* 3, 237 (1985).
- Gell, P.G.H. & Benacerraf, B. *Immunology* 2, 64 (1959).
- Ziegler, H.K. & Unanue, E.R. *Proc. natn. Acad. Sci. U.S.A.* 79, 175 (1982).
- Benacerraf, B. *J. Immun.* 120, 1809 (1978).
- Shimonevitz, R., Kappler, J., Marrack, P. & Grey, H. *J. exp. Med.* 158, 303 (1983).
- Schwartz, R.H. *Nature* 317, 222 (1984).
- Germain, R.N. & Mollison, E. *A. Rev. Immun.* 4, 281 (1986).
- Morrison, L.A., Lukacher, A.J., Lippman, V.L., Fan, D.P. & Braciale, T.G. *J. p. Med.* 163, 303 (1985).
- Cresswell, P. *Proc. natn. Acad. Sci. U.S.A.* 82, 8158 (1985).
- Willingham, M.C. & Pastan, I. *Cell* 21, 7 (1980).
- Pincus, M.R., Gerwitz, F., Schwartz, R.H. & Braciale, H.A. *Proc. natn. Acad. Sci. U.S.A.* 80, 3297 (1983).
- DeLisi, C. & Berzofsky, J.A. *Proc. natn. Acad. Sci. U.S.A.* 82, 7048 (1985).

Ronald N. Germain is in the Laboratory of Immunology, National Institutes of Health, Bethesda, Maryland 20892, USA.

Radionuclide deposition from the Chernobyl cloud

SIR—The accident at the Chernobyl nuclear power station in the Soviet Union created a cloud of radioactivity that affected a surprisingly large area of Europe. This letter concentrates on the main passage of this cloud over the United Kingdom and presents some results of preliminary investigations into the collected data on radioactivity levels in the air, in rain water, on grass and in milk, and attempts to link the time variations of these levels to meteorological conditions. In particular, we report that high levels measured on grass are linked to heavy rainfalls associated with thunderstorms embedded in the radioactive cloud.

The accident is believed to have begun at 01:23 local time on Saturday 26 April (20:23 GMT on Friday 25 April) and to have ended on the following Wednesday or Thursday. By the Monday, countries throughout Europe had been alerted and possible trajectories of the plume were being estimated, based on standard meteorological data. In many cases, analyses and forecasts from numerical weather prediction models were being used. Trajectories calculated at the Meteorological Office early in the week indicated a possible risk to the United Kingdom, and this possibility was strengthened when, on Wednesday enhanced radiation levels were reported in northern Italy, indicating a first arrival in the United Kingdom on Friday 2 May.

Since then, more thorough analyses of the cloud movement have been carried out at the Meteorological Office and at many other centres throughout Europe, and Fig. 1 shows the most probable path of the cloud that affected the United Kingdom. This part of the cloud appeared to leave Chernobyl at ~12.00 GMT on 26 April, ~16 h after the start of the accident. The cloud moved rather slowly within the lower layers of the atmosphere towards the Baltic, and then on the Monday a finger moved more quickly southwestwards over Poland, Czechoslovakia, Austria and the Alpine area, as a result of a growing zone of high pressure running across the Low Countries from south-west to north-west and a depression over the Adriatic. When this high moved away on Thursday towards the north-west, the cloud turned northwestwards over France towards the United Kingdom.

The frequency of environmental monitoring at a number of stations was increased, and the National Radiological Protection Board acted as a focal point for collating and interpreting these data^{1,2}. Reports of elevated radiation levels in southern England were received before midday on Friday 2 May and by late evening the cloud had reached the north

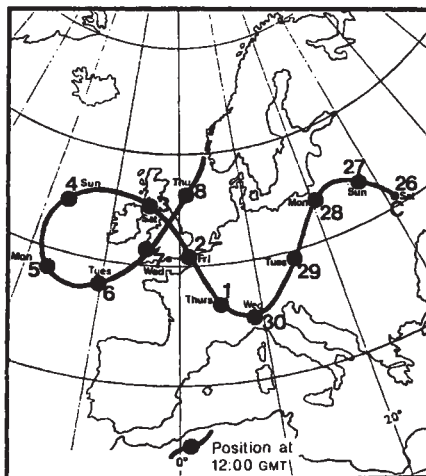


Fig. 1 Estimated path of the Chernobyl radioactive cloud which ultimately crossed the British Isles (26 April to 8 May 1986).

of England. The peak levels of γ -emitting radionuclides in the cloud were in the ranges shown in Table 1, and air concentrations in this range were observed later in Scotland and other parts of the British Isles, during the period 3–5 May. From this monitoring data and the data collected at the Meteorological Office we can derive a picture of the passage of the cloud, and Fig. 2a,b shows its position at 03:00 and 18:00 GMT on Saturday, 3 May. The cloud had entered Britain across the Kent–Sussex coast at ~05:00 GMT on the Friday morning, and moved northwards until by the end of the day most of Britain was covered except for Cornwall, Devon, the western extremities of Wales and north-west Scotland. Generally Friday was dry and sunny, the air being of continental origin, although a few showers were experienced in the Solway and Clyde areas late in the day.

On Saturday some old thunderstorms moving up from France were revitalized as colder air aloft was advected over the warm continental air, owing to a rapidly deepening low-pressure area to the southwest of the country. These storms first affected the south-east in the early hours, and then moved over East Anglia and out into the North Sea. New storms developed rapidly in the morning over central England. These moved slowly northwards to Cumbria and south-west Scotland, feeding off the warm air in the lowest kilo-

metre of the atmosphere containing radionuclides from Chernobyl, and causing the cloud to shrink in area. By early Sunday morning the cloud was affecting Scotland alone, and by 09:00 GMT it had been swept westwards from the Outer Hebrides, only to return, at much lower concentrations the following Wednesday and Thursday, having been swept right around the large low to the west. There is some evidence for a separate small part of the cloud having affected the United Kingdom on Sunday, when it seems the winds turned towards the south-east for a couple of hours, allowing a remnant from France to move northwards along the extreme east coast onto the far north of Scotland, where local rain gave high levels of radioactivity.

Radionuclides washed out by the rain produced elevated γ -ray backgrounds of up to $0.6 \mu\text{Sv h}^{-1}$ (about three times the normal level) in areas where it rained heavily, and this rain-out effect could be observed directly by measuring deposition densities on grass, as has been reported¹. The effect was reflected more directly in concentrations in rain water (Table 1), which varied from 10 to $1 \times 10^4 \text{ Bq l}^{-1}$ for ^{131}I . The highest levels were those reported from Scotland², which were very close to the derived emergency reference level (DERL) for ^{131}I in drinking water³. The DERLs are based on an effective dose equivalent of 5 mSv or an organ dose of 50 mSv, which ever is the more restrictive, corresponding to the annual dose limits recommended by the International Commission on Radiological Protection for exposure of the public over a few years⁴. The levels in rain water were followed by elevated levels in foodstuffs such as vegetables and cows' milk, but levels of ^{131}I in these foodstuffs did not exceed DERLs in any part of the country. The incorporation of radionuclides into animal tissues has been monitored by the Ministry of Agriculture, Fisheries and Food, which has issued data on meat, milk, vegetables and other foodstuffs⁵.

It is possible to use food monitoring data to infer deposition in areas of the country that could not be directly monitored, and match this with data on rainfall from the Meteorological Office. For example, subsequent monitoring of

Table 1 Ranges of peak air and rain-water concentrations of radionuclides observed in the British Isles during 2–5 May 1986

Radionuclide	Air concentrations (Bq m^{-3})	Rain-water concentrations (Bq l^{-1})
^{131}I	1–10	10–10,000
^{132}Te – ^{132}I	1–10	10–10,000
^{103}Ru	0.5–5	20–2,000
^{106}Ru	0.3–3	10–1,000
^{134}Cs	0.3–3	10–1,000
^{137}Cs	0.5–5	20–2,000
^{140}Ba – ^{140}La	0.3–3	10–1,000

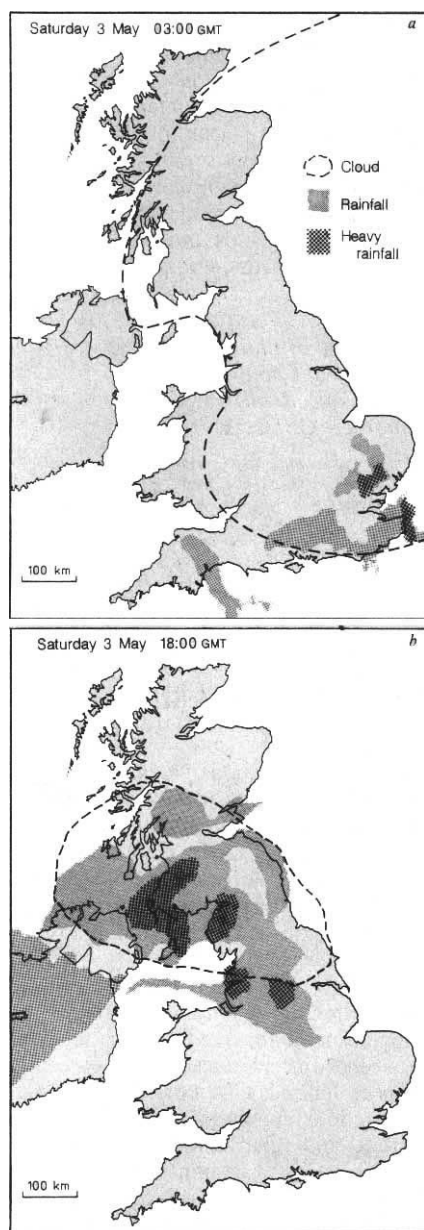


Fig 2 Chernobyl cloud and associated rainfall on 3 May 1986.

levels of ^{131}I in cows' milk² and sheep's milk³ from the south-east, East Anglia and the Midlands confirm that some wet deposition occurred in parts of southern England during the early hours of Saturday, 3 May (see Fig. 2), although the rainfall was not so prolonged as that observed later in parts of north Wales, north-west England and Scotland.

F.B. SMITH

The Meteorological Office,
Bracknell RG12 2SZ, UK

M.J. CLARK

National Radiological Protection Board,
Chilton, Didcot OX11 0RQ, UK

1. Fry, F.A., Clarke, R.H. & O'Riordan, M.C. *Nature* **321**, 193-95 (1986).

2. Levels of Radioactivity in the UK from the Accident at Chernobyl (HMSO, London 1986).

3. NRPB DL10 (HMSO, London 1986).

4. International Commission on Radiological Protection, *Ann. ICRP* **14** (2), (1984).

5. Ministry of Agriculture, Fisheries and Food Press Release, 30 May (1986).

How many reactor accidents will there be?

SIR—The designers of nuclear reactors claim that the chance of a reactor accident should be less than one in 10^6 per reactor year. The US Rasmussen report¹ and the German safety study² found that there is a fifty-fifty chance that such accidents could occur every 23,000 and 10,000 reactor-years respectively. Two or three serious accidents suggest the standard has not been met in practice, and that the method of 'technical risk assessment' used to calculate such probabilities must be replaced by risk assessments which use data from operating experience.

At the end of 1985, according to the IAEA data file³, there were 374 reactors in operation in the world with a total operating experience of 3,831 years. We assume that until the end of May 1986, the cumulative operating experience had been 4,000 years. There had been two major accidents—Three Mile Island and Chernobyl—during this operational period.

We do not count the Windscale of 1957 because it was a military reactor. We think that the reactor designers would claim that the safety systems of early reactors were not as good as the present ones. The British reactor designers were under pressure at that time. Military commitment to build several plutonium producing reactors came at a time when the civilian nuclear programme was emerging. The nuclear industry was expanding so quickly that there was an acute shortage of experienced scientists and engineers.

We assume that the accidents are a Poisson process with intensity r . The probability of finding two accidents in an interval of length $T = 4,000$ reactor-years is then Poisson distributed:

$$P(2, rT) = r^2 T^2 e^{-rT}/2 \quad (1)$$

If we want this to be a probability density, we have to normalize it. The normalization being:

$$N(T) = \int_0^\infty dr P(2, rT) = 1/T \quad (2)$$

The density is then

$$P_r = r^2 T^2 e^{-rT}/2 \quad (3)$$

The numerical value of P_r is not a probability. However, we can think of $P_r dr$ as the element of probability—for example, the probability that r lies between r_1 (0.0004) and r_2 (0.006) is obtained by integrating equation (3) between r_1 and r_2 and is found to be 0.21. This indicates that the probability that one significant accident will occur between 1,667 and 2,500 reactor years is 21%.

Another way of dealing with this density is to take into consideration that the

probability of the occurrence of at least one accident in a Poisson process of intensity r in a time interval t is

$$P(1, rt) = 1 - e^{-rt} \quad (4)$$

Since r is unknown, we have to integrate equation (4) overall r with the 'empirical' density P_r as a weighting factor³, that is:

$$P(1, t) = \int_0^\infty dr (1 - e^{-rt}) r^2 T^2 e^{-rT}/2 \\ = 1 - (1 + t/T)^{-3} \quad (5)$$

Equation (5) gives the probability of having at least one accident during the next t years. The probability that a reactor accident *could* happen during the next decade ($t = 374$ reactors $\times 10$ years = 3,740 reactor-years) is 86%.

There is no circularity in the arguments used above. There must be some input data in any model. Our input is two accidents in 4,000 reactor-years. Of course, we know that two accidents are not 'representative'. With 374 reactors in operation, we would naively expect one accident every 5.4 years. But our model assigns a 70% probability that one accident *could* happen in the next 5.4 years, not a certainty. Similarly the probability of having one accident every two decades is more than 95%.

It is true that events with high probability need not necessarily materialize. But how can we reconcile these findings with the claims of the reactor designers that the chance of a reactor accident should be less than 10^{-6} per reactor-year? The reactor designers calculate such probabilities with

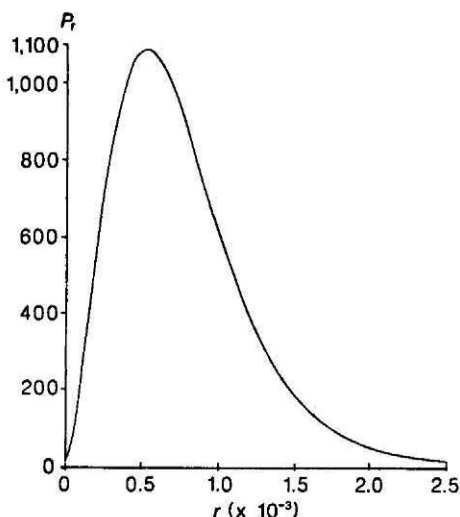


Fig. 1 P_r gives the likelihood of a significant accident given the total operating experience of 4,000 reactor-years and two significant accidents. The number of significant accidents per reactor year is measured by r . The probability that this number lies between r_1 and r_2 accidents per reactor year is given by the area under the curve between r_1 and r_2 , that is, by integrating equation (1) between r_1 and r_2 .

the help of a decision tree. The likelihood of the failure of a reactor component or safety system is assessed and its consequences on other components or subsystems are estimated. These results in their turn are inputs for a chain of calculations leading to probabilistic risk assessment used by the reactor designers. Since the number of vulnerable components is rather high, this procedure cannot be rigorous. Our view is that this method should be replaced by risk assessment using the observed data.

We thank Professor K.E. Eriksson of Göteborg and Dr K. Hildenbrand of Bonn for helpful discussion and correspondence.

S. ISLAM

International Study Group on Self-organization
Bohlweg 44,
3300 Braunschweig,
FRG

K. LINDGREN

*Physical Resource Theory Group,
Chalmers University of Technology,
University of Göteborg,
41296 Göteborg,
Sweden*

1. *Reactor Safety Study* (Nuclear Regulatory Commission, Washington DC, 1975).
2. *Die deutsche Risikostudie* (Federal Ministry of Research and Technology, Bonn, 1979).
3. *IAEA Bull.* 28(1), 71 (1986).

¹³¹I in ruminant thyroids after nuclear releases

SIR—The substantial yield of ¹³¹I from nuclear fission, its energetic β and γ emission and its concentration in the thyroid gland have led to the recognition of this nuclide as a significant nuclear hazard^{1,3}. In the late 1950s the Medical Research Council funded a study of ¹³¹I accumulation in the thyroid glands of sheep from fallout arising from nuclear weapons tests. This study was in progress during the accident at Windscale in October 1957 and during the Northern Hemisphere weapons testing of October–December 1958 in which 29 (17 Soviet and 12 American) weapons were detonated¹. British weapons at that time were detonated in the Southern Hemisphere, having little effect on fallout in the United Kingdom².

It was therefore of interest to measure ¹³¹I accumulation in sheep thyroid glands following pasture contamination with fallout from the Chernobyl reactor, to obtain a comparison with the earlier measurements. Thyroid glands were collected from fatstock from 20 May to 17 June. At this time sheep were at pasture and hence directly ingesting any nuclides on the grass. Because of the late spring, fat cattle were generally not at pasture at the time of maximum contamination, on 2 May⁵.

Figure 1 shows thyroid ¹³¹I radioactivity from sheep grazing in Scotland during the autumn of 1958⁴ and in East Anglia during

the spring of 1986. It can be seen that thyroid radioactivity following the Chernobyl accident was marginally higher than that resulting from the nuclear weapons testing during the autumn of 1958. By contrast, the highest sheep thyroid ¹³¹I measured in the maximum deposition area near Windscale 31 days after the reactor accident was 7.4×10^4 Bq g⁻¹, 300 times higher than the radioactivity measured after Chernobyl. In the sheep thyroid glands recently measured, the radioactivity had a half-life on storage of 8.05 days, showing that ¹³¹I alone was present. Comparison of radioactivities on different dates of collection showed a biological half-life of about 6 days.

Cattle thyroid glands are less useful than those of sheep for monitoring fallout, because of the widespread practice of feeding beef cattle indoors on stored fodder for much of the year. Thyroid ¹³¹I in stall-fed sheep following a period of radioactive fallout from nuclear weapons testing by the United States showed only 0.7% of the radioactivity in grazing sheep⁴; a limited number of human thyroids also showed ~1% of the radioactivity of sheep¹. Five cattle thyroids collected recently in East Anglia showed a mean of 0.97 ± 0.13 Bq g⁻¹, compared with 99 ± 5 Bq g⁻¹ for sheep killed the same day. One other cow, however, had 61.7 Bq g⁻¹ ¹³¹I, which must have resulted from grazing on contaminated pasture.

Data already published⁵ showed human thyroid radioactivities after Chernobyl of 2–33 Bq, the highest concentration being ~7 Bq g⁻¹, in a child, while adults showed a maximum radioactivity of 1.7 Bq g⁻¹. Sheep thyroid radioactivity on the same date measured 243 ± 48 Bq g⁻¹ ($n=15$), with a range of 39–703 Bq g⁻¹.

¹³¹I from a reactor accident is capable of contaminating pasture for a considerable distance downwind. Most human inges-

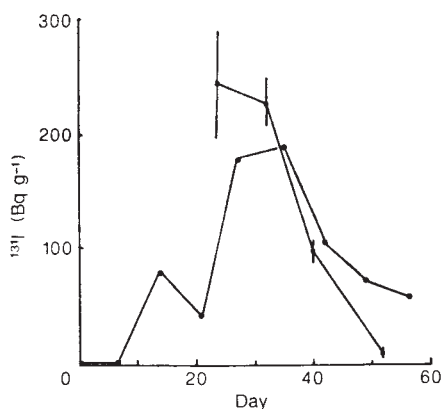


Fig. 1 Sheep thyroid radioactivity following the reactor accident at Chernobyl (O) and the nuclear weapons testing of the autumn of 1958 (●). Day 0 has been taken as the day of the Chernobyl accident (26 April 1986) or the peak of nuclear weapons testing (3 November 1958). ¹³¹I activity is given in Bq g⁻¹ wet weight of gland ($\pm$ standard error of mean).

tion of ¹³¹I derived from fallout comes through milk, as the mammary gland effectively concentrates dietary iodine into milk⁶. To minimize the human ingestion of ¹³¹I, a reasonable precaution would be to feed dairy cattle stored fodder while pastures are measurably contaminated.

I thank Mr Martin Collins for collection of the thyroid glands and the University Pathology Department for the use of their γ -counter.

IAN R. FALCONER

*Department of Biochemistry,
University of Cambridge,
Tennis Court Road,
Cambridge CB2 1QW, UK*

1. Van Middlesworth, L. *Science* **123**, 982 (1956).
2. Marston, H.R. *Aust. J. biol. Sci.* **11**, 382 (1958).
3. Maycock, G. & Vennart, J. *Nature* **182**, 1545–1547 (1958).
4. Robertson, H.A. & Falconer, I.R. *Nature* **184**, 1699–1703 (1959).
5. Hill, C.R., Adam, J., Anderson, W., Ott, R.J. & Sowby, F.D. *Nature* **321**, 655–656 (1986).
6. Robertson, H.A. & Falconer, I.R. *J. Endocr.* **22**, 133–142 (1961).

Fault striations slip from sight?

SIR—Communication of ideas begins with agreement on the meaning of terms. Unfortunately, in discussions on mesofaults and their surface textures, terms are endowed with different meanings with alarming ease thus giving rise to considerable confusion. A glaring example is the term 'slickenside'.

Linear structures on fault planes (grooves or fibrous grains) represent the orientation of the relative displacement between adjacent fault blocks. Recently many techniques have been developed to determine from slip orientation data the attitude and relative magnitudes of the principal stresses responsible for the deformation. Such data are of paramount importance in the study of neotectonics relating recent plate motions to deformation within continents. The problem is that some authors^{1,2} use slickenside to refer to planar structures whereas others² use it to describe a lineation.

Etymologically slickenside means smooth plane, which is how it was first defined in 1822. Surely this should remain as the agreed meaning. The associated linear structures are best termed striations; qualifying words, such as fibre and scratch, are adequate to distinguish accretionary crystal growths from abrasion structures³.

CLIVE A. BOULTER

*Department of Geology,
University of Nottingham,
University Park,
Nottingham NG7 2RD, UK*

1. Fleuty, M.J. *Geol. Mag.* **112**, 319–322 (1975).
2. Dennis, J.G. *Am. Ass. Petrol. Geol. Mem.* No. 7 (1967).
3. Peltzer, G., Tapponnier, P., Zhang Zhitao & Xu Zhi Qian. *Nature* **317**, 500–505 (1985).
4. Sporli, K.B. & Anderson, H.J. *N. Z. J. Geol. Geophys.* **23**, 155–166 (1980).

Star on the make?

Colin Patterson

The Nemesis Affair: A Story of the Death of Dinosaurs and the Ways of Science.
By David M. Raup. Norton: 1986. Pp. 220. \$15.95, £12.

"I HAVE never seen Francis Crick in a modest mood" is one way to start a book; "This is the story of an emerging scientific theory about the extinction of dinosaurs and other prehistoric life forms" is another. With the first sentence of *The Double Helix* we know just where we are — the academician is out of school, the bag is open and the cat is flying, together with whatever else was in there. The opening of *The Nemesis Affair* does not set the scene so plainly. Is the academician in or out of school, and if in, what class are we with?

Of course it is unfair to compare David Raup's book with Watson's, but Raup is used to the ways of science and knows they are not paved with fairness. It is not often that a leading professional tells the tale while still close to the events — Raup calls it "a view from the trenches" — and the Watson/Raup comparison does bring out aspects of the ways of science. For one thing, the pace is killing today: there was a gap of 15 years between the events Watson wrote of and his book, whereas Raup hits the streets only two years after the birth of Nemesis, the name given in *Nature* of 19 April 1984 to a possible companion star to the Sun, producing periodic comet showers and mass extinctions.

Nemesis was born as one of several possible causes for a possible series of events. In February 1984 (*Proceedings of the National Academy of Sciences*, **81**, 801–805), Raup and Jack Sepkoski published a statistical analysis of Sepkoski's data on the first and last appearance of each marine family in the fossil record. They concluded that extinction is non-random, with a 26 Myr periodicity; one of their eight best-fitting mass extinctions is at the Cretaceous/Tertiary boundary, hot property since 1980 because of the iridium spike which may imply extraterrestrial impact. Six weeks after the Raup and Sepkoski paper in *PNAS* came the 19th April *Nature* containing an editorial by John Maddox, a "News and Views" piece by Tony Hallam, and a block of five papers, all bearing on extraterrestrial causes of periodic mass extinction. (Turning back to *Nature* in 1953, the Watson and Crick paper didn't merit mention in "News and Views", which then read rather like the Court Circular.) Maddox's editorial commented (Raup calls it "mild wrist slapping") on the practice of distributing preprints, which in this case resulted in the

possibility of the *Nature* papers appearing *before* the Raup and Sepkoski data they purported to explain. John Maddox saw this as "a kind of nonsense", but at least it is nonsense with a limit. The pace can get no hotter once the hound and hare coincide.

Since 1984, it is hard to pick up a journal in the fields of geology and evolutionary biology which doesn't contain a paper, review or meeting report touching on mass extinctions. DNA was slower to get up steam, but once it got going it was unstoppable, and we all know what has been achieved. Can we expect a similar profoundly important research programme from Nemesis *et al*? I think not, and Raup gives no real sign that he thinks otherwise. He ends his book with some comments on the pecking order from "hard" to "soft" science, placing molecular biology at the hard end and palaeontology (his own field) at the soft. We can see now that Watson and Crick's achievement was to begin to move parts of biology from the historical (soft) towards the physical (hard) end of the spectrum — to put some determinism among the contingency. Periodicity in mass extinctions, and extraterrestrial explanations for it, might seem to promise the same for parts of palaeontology. But, so far, the controversies over periodicity and the less general question of extraterrestrial impact as the cause behind particular extinctions exemplify the field of sociology (which Raup puts right at the bottom of the pecking order) as much as hard science. Recalling Raup's "view from the trenches", those involved seem first instinctively to have either volunteered for the neocatastrophist revolutionary force or

rallied to the flag of Lyellian uniformitarianism, and then started looking around for ammunition. (The latest skirmish in *Nature* was "Matters Arising" in May this year (**321**, 533–536).)

Raup has some level-headed comment on the prevalence of prejudice and preconception in his chapter on "Belief Systems in Science", but ultimately the Raup and Sepkoski periodicity thesis (and so the possibility of Nemesis) rests on the quality of Sepkoski's taxonomic data. That is where I believe the thesis is weakest, but the attack will necessarily be slow and piecemeal. *The Double Helix* was a success story, whereas Raup's book may turn out to be (as he acknowledges) just the happier half of a chronicle of failure. If so, he may find a benefit in avoiding "saganization", the word he coins for the (unmerited) fall in professional esteem that accompanies the rise in visibility of any scientist taken up by the media.

Like *The Double Helix*, Raup's book can be read at a sitting. What gripped the reader in Watson's story was the personalities, the emotions, the indiscretions. Perhaps because he is still so close to the events and to the people, Raup has left all that out; his colleagues or opponents are characterised as no more than "brilliant", "prominent" or "hardworking". No one will take offence at the book, but the lack of emotion or indiscretion means that the narrative and the writing have to do a lot of work. On the whole, they stand up to the strain; the pace of the narrative picks up in the second half of the book once the scene has been set, and the writing is good popularization, though occasionally slipping a bit too far down market. The first words of "what is known as the Upper Eocene" or "what are called the Middle and Upper Jurassic" are no sop to anyone, and the sentence that Raup and I (unfairly) begin with does not set the true tone of the book. □

Colin Patterson is a vertebrate palaeontologist at the British Museum (Natural History), London SW75BD, UK.



When they stayed away from Paris

Robert Fox

Science in the Provinces: Scientific Communities and Provincial Leadership in France, 1860-1930. By Mary Jo Nye, *University of California Press*; 1986. Pp.328. \$39.95, £33.95.

OVER the past decade, the community and institutions of French science in the nineteenth and early twentieth centuries have been subjected to an unprecedented onslaught by historians. Oddly, the trail has been blazed by scholars working in Britain and North America, and it is they above all who have fashioned the developing consensus to which Mary Jo Nye's book makes a notable contribution. In the new orthodoxy, two points have emerged strongly. The first is that in the half century or so before the First World War the majority of French *savants* abandoned their traditional preoccupation with abstraction and pure knowledge and became increasingly concerned, even besotted, with the applications of science. The second is that the scientific life of the provinces deserves to be treated quite as seriously as that of the capital and that it is quite misleading (though understandably convenient) to equate an assessment of Parisian science with that of French science as a whole.

Both of these points are developed to good effect in *Science in the Provinces*. The book is structured around a series of studies of the science faculties of Nancy, Grenoble, Toulouse, Lyons and Bordeaux, and it uses those studies to display the faculties as seats of scientific vitality and institutional innovation, with local industrial agricultural and commercial interests often serving as a source of problems and a spur to action. Strikingly, Nye's faculties all lie in a crescent sweeping from the Gironde, through the Midi and up into Lorraine. Here, it seems the provincial tradition in science developed with special vigour: certainly, it would be hard to think of professors in the faculties of the centre or the north-west of France — at Clermont-Ferrand, Poitiers or Rennes, for example — who would bear comparison with the undisputed "stars" who sustain Nye's case for each of her five centres.

By digressing liberally on the work of individuals, rather than attempting a comprehensive prosopography, Nye allows the content of the science she treats to come to the foreground, where it belongs. Not all the science has stood the test of time, to be sure: René Blondlot's false discovery of a new radiation — N-rays — in 1903 is a case in point. But by the time

Victor Grignard (left) and Paul Sabatier refused the call to Paris

Victor Grignard (of Lyons and then Nancy) and Paul Sabatier (of Toulouse) shared the Nobel prize for chemistry in 1912, the provincial faculties of science had to be taken seriously, the more so as certain professors (including Grignard and Sabatier) ostentatiously refused the call to a Parisian chair. Their resistance was understandable: good facilities, local adulation and a sense of contributing to the regional economy were the rewards on offer and, for about 20 years from 1900, they sufficed to reduce the flow of talent to the capital. In the longer run, however, what seems to have been the natural order of things reasserted itself. By the 1930s, the provincial faculties had become once again the poor relations of a crippling impoverished university system, and even Grignard had lost his old assertiveness and hope that the imbalance between Paris

and the provinces might be removed.

Nye's discussion of Grignard adds poignancy to the story of a valiant but doomed campaign for diversity. It is a story well worth telling and with some bearing on present educational debates not only in France but also in Britain. It suggests, at least to a reader of gloomy disposition, that although centres of excellence can be fashioned on the margins of the traditional academic world, they are intrinsically fragile creations condemned to a thorny path and, in all probability, an ephemeral existence. □

Robert Fox is British Academy Reader in the Humanities at the University of Lancaster, Lancaster LA1 4YG. From September, he will be Director of the Centre de Recherche en Histoire des Sciences et de l'Industrie at the Cité des Sciences et de l'Industrie, 75930 Paris Cedex 19, France.

Stones in the service of society

D.V. Ager

The Stones of Britain: Landscapes and Monuments, Quarries and Cathedrals. By Richard Muir. *Michael Joseph*; 1968. Pp.288. £15.95.

WHEN we were in the United States and our daughter was only one year old, she greatly impressed the locals by referring to the erratic blocks of the glaciated Mid-West as "rocks" rather than "stones". That came of growing up in a geological family. Geologists do not study "stones" unless they be of the "precious" or "semi-precious" variety. The only other exception is building "stones", which are the main subject of Richard Muir's book.

Muir is a geographer and tells us right at the beginning that he is not writing about geology because "geology is a technical

subject". So this is a book for the architect and the aesthete rather than the scientist. It is a book about man in Britain using the rocks beneath his feet, first for implements, then for monuments and finally for buildings.

The first part deals with "stone" in the landscape. This Muir divides into three: "Hard Rock Landscapes", which means chiefly the granites of Cornwall and the Ordovician volcanics of the Lake District; "Sandstone Scenery", which curiously includes the brick-making clay-pits; and "Landscapes of Limestone", which includes the chalk but hardly the flints which that lithology contains.

Flints come into their own in the next part of the book, on the use of stone by prehistoric communities. Here are accounts of stone tools (which, it is worth recalling, served man for nearly 99 per cent of his history and of which there is a particularly good record in Britain), cave dwellings (somewhat inappropriately) and then the first stone huts, starting with Skara Brae in the Orkneys. I have a theory

that the oldest buildings in Europe — in the Orkneys and in Malta — are still there simply because of the ease with which the local stone could be worked so that it was not worth the trouble of pulling the huts down again to build more up-to-date residences. Muir then moves on to burial chambers and stone monuments. He makes some rather doubtful remarks about Stonehenge but does not mention the unlikely theory (albeit sanctified in the Geological Museum) that the "blue stones" were brought from Wales by ice rather than by human muscle.

The third and most interesting part of the book deals with what the author calls "The Rise and Fall of Building Stone" in the great mediaeval cathedrals and castles. The biggest problem was the cost of transport before the coming of canals and railways, when the movement of heavy materials was largely restricted to natural waterways. The local stone was mostly Jurassic limestone or Devonian sandstone, though other local rocks were obviously used when convenient. Muir describes quarrying methods, and the mason's art in relation to changing architectural styles, and relates economic and social history to the remarkable phenomenon of great stone buildings which only ended (as I see it) with gunpowder, Ann Boleyn and the Reformation.

There is much less in the book about domestic architecture, the triumph of which was surely the cottages and walls of the Cotswolds with the "sunlight" in their local Middle Jurassic limestones, and even less on post-mediaeval building in stone. The use of Permian Magnesian Limestone in the Houses of Parliament is mentioned, though not the disastrous effects of the acid London atmosphere in turning them into Epsom Salts!

Muir talks disparagingly of brick (also apparently a form of "stone"), but does not, in my view, say enough about the uppermost Jurassic Portland Stone which Wren used to rebuild church and state in London after the great fire of 1666 and which then spread throughout southern Britain as the pale face of bureaucracy. Thus, characteristically, it is used in Swansea for the Guildhall, whilst the prison is built more prosaically of the local Pennant Sandstone. Banks, on the other hand, seem to depend on the Aberdeen granites to give them a proper air of dependability, while Cornish granites provide our hard-wearing kerb-stones.

For its intended audience this is a fascinating book, well-written and easy to read. Personally, I would have liked to have had also a simpler guide, little more than: "Salisbury Cathedral — Chilmark Stone" and then where Chilmark is and the geological age of its building stone. □

Derek Ager is Research Professor in the Department of Geology, University College of Swansea, Swansea, SA2 8PP, UK.

Fascinating family

Frances C. James

Blackbirds of the Americas. By Gordon Orians. *University of Washington Press: 1985. Pp.163. \$24.95, £18.*

WHEN American ornithologists use the term "blackbird", they mean a member of the subfamily Icterinae, a group confined to the Western Hemisphere. Among the 94 species are not only the familiar red-winged blackbird, the brown-headed cowbird and the common grackle, whose immense roosts are a nuisance to farmers, but also many colourful species of New World orioles and meadowlarks, and the tropical oropendolas and caciques. The lineage is as diverse in social behaviour and breeding biology as it is in physical appearance and distribution.

Gordon Orians' fascination with blackbirds has taken him from Canada to Argentina, from marshes to deserts, from savannas to tropical forests. In this attractive volume, he introduces American blackbirds to the layman, the bird-watcher who has mastered the field guide and wants to learn more. His informal

writing style, in combination with the many striking drawings by Tony Angell, makes the book seem friendly, even though it is densely packed with information. And the comparative approach, *à la* David Lack, leads to new perspectives on the subject matter.

In an attempt to touch more generally on what evolutionary biologists do, Orians describes his own experiments and those of others. He discusses applications of optimal foraging theory and kin selection theory to interpretations of behavioural data, and confesses that the work was not designed to test these constructs but merely to use them to organize information. Here both the strengths and weaknesses of a controversial area of biology are revealed. The sceptical layman may well decide that biologists are still a long way from the goal of understanding mechanisms of adaptation.

The book is dressy, rich in detail and yet easy to read. References and appendices are included for the reader who wants to go further. Orians has fully achieved his objectives, perhaps even more than he intended. □

Frances C. James is a Professor in the Department of Biological Science, The Florida State University, Tallahassee, Florida 32306, USA.

For the future

Lynton K. Caldwell

World Resources 1986: An Assessment of the Resource Base that Supports the Global Economy. World Resources Institute and International Institute for Environment and Development. *Basic Books/Harper & Row:1986. Pp.353. Hbk \$32.95, £22.95; pbk \$16.95, £11.95.*

THIS reference source is the first volume of a projected annual series. Aspects of ten resource-environment topics are reviewed, with data tables for each of them and also for selected materials, minerals and basic economic indicators. Subsequent volumes will cover other aspects of the subject and a new one, on biogeochemical cycles, is planned for 1987.

Among the topics covered are population; human settlements; food and agriculture; forests and rangelands; wildlife and habitat; energy; freshwater; oceans and coasts; atmosphere and climate; and policies and institutions. The editors acknowledge the difficulty of organizing a volume in which all the material is interrelated, but while some division by topic was unavoidable and artificial the reader is assisted by a detailed table of contents, an index and frequent cross-referencing.

A further difficulty in compiling a book such as this is the frequent inadequacy of

the information available. The editors found "gaping holes in basic knowledge", and the prospects for closing these gaps are often diminished by lack of reliable indicative criteria and of comparability of the data. One of the main objectives of the *World Resources* series is to help remedy these deficiencies, and to provide assessments of the global resource base that can be relied upon in the formulation of national and international policies for resources and environment.

World Resources 1986 is an impressive compilation, well-designed for ready reference. Some minor textual errors are inevitable in a production of this kind, but I found no misleading errors of substance. It is not, however, the only source of information on world resources and environment. Since 1984 the Worldwatch Institute, under the direction of Lester R. Brown, has published an annual *State of the World* report — but although the two publications deal with many of the same issues they are complementary rather than competitive. Both are predicated on the proposition that the world environment is rapidly deteriorating under human mismanagement, but can be saved if timely action is taken. Both avoid the "gloom and doom" approach to the predicament of mankind. In 1984 the World Resources Institute sponsored The Global Possible Conference, principal papers from which were published in 1985 by Yale University Press. Similarly, the Worldwatch volume is advocacy for an

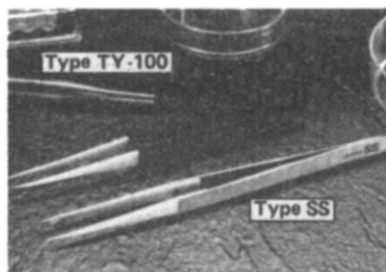
ecologically sustainable society which its editors and authors believe to be possible.

Having emphasized the virtues of these reports, I should point out what they do not offer. They do not provide a comprehensive survey of scholarly and scientific research on the topics covered (nor should this be expected). And, although they are policy-orientated, I believe that few social scientists would regard their views on such matters as adequate or indeed as realistic.

World Resources 1986 is largely a straightforward and factual account. I have no quarrel with the conclusions reached in the treatment of policies and institutions, but the report as a whole does not address the question of the human will to achieve the "global possible". Neither, unfortunately have social and behavioural scientists seriously addressed this question although some have begun to ask it. The "global probable" would have more to do with values, assumptions, ideologies and behavioural patterns than with ecosystems or energy production. An assessment of human prospects for realizing the possible could be very useful in developing strategies for a sustainable future. □

Lynton K. Caldwell is Arthur F. Bentley Professor of Political Science Emeritus, *Advanced Studies in Science, Technology, and Public Policy*, Indiana University, Bloomington, Indiana 47405, U.S.A.

NEW CERAMICS PRECISION: FUNA-TWEEZERS



"FUNA-TWEEZERS" can eliminate all demerits of stainless or plastic tweezers. The stainless stems of Type SS and the Derlin stems with glass reinforcing agent of Type TY-100 give appropriate elasticity.

These show excellent features such as high anti-corrosiveness, heat insulation, high flame resistance, etc. These are also free from plastic deformation, metallic contamination and magnetization, etc.

These tweezers are contributing to the industries and research such as Biotechnology, Electronics, etc.

FUNAKOSHI PHARMACEUTICAL CO., LTD.
Science Business Div. Research Material Sect.
2-3 Surugadai, Kanda, Chiyoda-ku, Tokyo, Japan
Telephone: Tokyo 03-293-2352
Telefax: 81-3-295-5545
Telex: J28489 FUNA

Energy for life

Owen T.G. Jones

The Vital Force: A Study of Bioenergetics.

By Franklin M. Harold. W.H. Freeman: 1986. Pp.577. Hbk \$37.95, £19.95.

BIOENERGETICS attempts to explain how cells generate useful energy and how this conserved energy is used in the enormous range of activities that can be defined as work. As these activities include the capacity to respond to changes in the environment, to move, to grow and to reproduce, it is apparent that a comprehensive account of bioenergetics could encompass much of biochemistry and cell biology. Dr Franklin Harold has accepted the challenge in *The Vital Force*. He not only covers well the expected topics of oxidative phosphorylation, photophosphorylation, ion transport, active transport processes and muscle contraction, but also includes the motility of bacterial and animal cells and the involvement of ion fluxes in the transduction of signals from the cell membrane to the cytosol and the nucleus.

It is, of course, ambitious of one author to range over topics which are so extensive and which are developing rapidly, but his attention to the bioenergetic aspects of these complex events permits him to disentangle common properties of apparently disparate processes. The significance of protein phosphorylation in regulating cellular activities and the ever-increasing evidence of the importance of changes in the concentration of intracellular Ca^{2+} and other ions in cell division and differentiation support Harold's wide-ranging coverage.

The book opens with a comprehensible and simplified account of the thermodynamics necessary to understand problems of bioenergetics. The author adopts the view that the major disputes of the chemiosmotic wars are now settled and that energy conservation is by mechanisms closely related to those proposed by Peter Mitchell. He does, however, give an account of the background to these disputes — such as whether energy coupling arose from protons extruded into the bulk phase or from protons localised within the anhydrous phase of the membrane — and also explains why there have been such a range of values for the ratio of protons translocated during electron flow to ATP produced. In following his explanations, the reader will gain an understanding of the thermodynamics of ion movement and ATP synthesis and why it is important. From this clear account the reader will also learn what a proton-motive Q cycle is why it is necessary for a satisfactory formulation of the chemiosmotic hypothesis, and the relative merits of redox loops

and conformational pumps as mechanisms to transport protons.

The author is particularly familiar with the energetic problems of microorganisms and he writes on these with authority. Bacteria can face alarming problems of maintaining osmotic pressure and cytoplasmic pH near neutrality whilst growing at a pH 2 or pH 11 or in salt concentrations around 5M. Their strategies for coping with life at these extremes are not fully understood but Harold does show how much information of very general significance can come from the study of relatively obscure organisms. Halobacteria, for example, have the simplest of proton pumps and the determination of its structure has been a great stimulus to thought. One minor criticism is that Harold's description of the reaction centre of photosynthetic bacteria does not include an account of its crystal structure which has so elegantly confirmed the arrangement of pigments and electron transport carriers which he describes. These had been predicted by bioenergeticists from their more usual techniques of measurement of flash-induced spectroscopic changes during the generation of light-driven transmembrane electric field. Such structural studies on complex protein assemblies are at last starting to answer some of the important questions about fundamental mechanisms in energy conservation.

The Vital Force is a substantial volume, in size and content. It will be a useful reference book for the specialist and will provide a good initiation for the more general reader. In the area of accepted bioenergetics its coverage is comprehensive and authoritative. On the topics Harold is now claiming for bioenergetics — the cytoskeleton the generation of intracellular signals and morphogenesis — he brings to our attention much important and exciting work. The measurement of the movement of Ca^{2+} ions is likely to provide as much employment for biochemists and cell biologists in the future as has the measurement of proton movements in the past. □

Owen T.G. Jones is Professor of Biochemistry, Medical School, University of Bristol, Bristol BS8 1TD, UK.

New Editions

- *Energy and the Atmosphere*, 2nd Edn. by I.M. Campbell (Wiley: hbk \$57, £34.95; pbk \$22.25, £13.50).
- *State of the World 1986*, by Lester Brown et al., Worldwatch Institute. (Norton: pbk \$10.50, £7.50).
- *Techniques in Bioproduction and Photosynthesis*, 2nd Edn. edited by J. Coombs et al. (Pergamon: hbk £24.50, \$32; pbk £12.75, \$17).
- *The Periodic Table of the Elements* (2nd Edn), by R.J. Puddephatt and P.K. Monaghan (Oxford UP: hbk £15, \$27.50; pbk £6.95, \$10.95).

Gene regulation by proteins acting nearby and at a distance

Mark Ptashne

Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

Transcription of genes can be controlled by regulatory proteins that bind to sites on the DNA either nearby or at a considerable distance. Recent experiments suggest a unified view of these apparently disparate types of gene regulation.

THERE are many examples in prokaryotes and eukaryotes of the following phenomenon: a protein bound at a specific DNA site influences the transcription of a gene hundreds or even thousands of base pairs away. For example, in eukaryotes, proteins bound at a site called an enhancer can turn on transcription of a distant gene. Regulatory proteins thus influence the activity of another protein (RNA polymerase) at a distance—how is this possible? Those who would tentatively throw all such examples into one basket and search for a common mechanism (and I am one) divide into distinct camps: the twisters, sliders, oozers and loopers.

My purpose is to review selectively recent experiments that support the loopers and to varying degrees weigh against the first three camps. The discussion unfolds within the context of our firm understanding of how one class of regulators, exemplified by the λ -phage repressor, recognizes DNA and turns transcription on or off. In that case proteins bind to adjacent sites on helical DNA and regulate transcription by touching one another, a matter I explain more fully below. The loopers imagine that proteins bound at widely separated sites act in the same way, contacting each other, with the intervening DNA looping or bending to allow the protein-protein interaction. According to this idea, it is the interaction between DNA-bound proteins, not the looping *per se*, that regulates gene expression.

In the course of this article I shall first consider the evidence for looping and then define and argue against the contrasting mechanisms. I shall consider experiments involving both gene transcription and site-specific recombination. The latter phenomenon (at certain stages) requires that proteins acting at separated sites on the same DNA molecule communicate with each other, and mechanisms for how that interaction occurs have been critically tested.

The case for looping

The first indication of looping came from studies of gene regulation in *Escherichia coli*. Adhya and co-workers^{1,2} showed that two operators separated by some 90 base pairs are required for efficient repression of the *gal* operon and suggested, as one possible mechanism, interaction between Gal repressors bound at these operators. Schleif and co-workers³ similarly discovered that complete repression of the *araBAD* operon depends upon an operator site to which the regulatory protein AraC binds, located some 200 base pairs upstream of the transcription start site. They deduced that the AraC protein bound to the upstream operator interacts with another protein bound near the transcriptional start site, the intervening DNA looping to allow the interaction. This conclusion followed from the remarkable observation that efficient repression was maintained when an integral number of helical turns of DNA was added or deleted between the operator and the gene, but not when the spacing was varied by half-integral numbers of turns. They reasoned that proteins separated by the relatively short distance of a few

hundred base pairs can interact by simple looping only if located on the same side of the helix; introduction of half-integral turns would place the proteins on opposite sides of the helix and the energy required both to bend and twist the DNA would be prohibitive. Schleif and colleagues^{4,5} have since presented additional evidence consistent with the idea of DNA looping.

The ideas discussed above have recently been extended to the case of a mammalian virus, in particular the early gene promoter of simian virus 40 (SV40). This promoter contains three adjacent sets of sequences that influence transcription: the 'TATA' region near the transcription start site; a middle segment of 60 base pairs called the 21-base-pair repeats; and further upstream the enhancer, covering some 150 base pairs. Efficient transcription evidently requires that proteins bind to each of these segments. Takahashi *et al.*⁶ report that the insertion of 5 or 15 base pairs between the enhancer and the middle segment decreases transcription *in vivo* more drastically than does insertion of 10 or 21 base pairs. Similar results have been found for insertions between the middle segment and TATA. The former result is particularly striking because the enhancer functions even when re-positioned hundreds of base pairs from the middle segment. A simple interpretation of these experiments is that proteins bound to the enhancer contact other proteins bound to the middle segment and that these in turn contact a protein bound at TATA. When the enhancer is re-positioned at a distance (and perhaps in the normal position as well) the DNA presumably loops to allow the interaction to take place.

Our most direct demonstrations of interaction between adjacent DNA-bound proteins and between separated proteins with concomitant DNA looping come from studies of the λ -phage repressor. Figure 1 and its legend summarizes several aspects of the structure and activities of this protein. Briefly, the repressor, which can activate as well as repress transcription, normally binds to two adjacent sites on the λ chromosome. A repressor dimer bound at the stronger of the two sites helps a second dimer bind to the weaker site. This cooperative binding, which depends upon contact between the proteins, increases the affinity of the weaker site some 10-fold. The repressor at the weaker site then directly contacts RNA polymerase and helps it to bind and begin transcription of the adjacent gene. At each operator site, repressor is bound along one face of the helix (for a review see ref. 7).

Thus, interacting λ repressors normally lie adjacent to each other on the DNA; our recent experiments show that they also interact when separated. Hochschild and I⁸ used the technique of 'DNase footprinting' to show that λ repressor binds cooperatively to operator sites separated by approximately integral numbers of turns (five, six and now seven). Just as when the sites are in their normal positions, binding to the strong site increases binding to the weak site some 10-fold. In contrast, cooperative binding to operators separated by, for example, 6.5 turns was observed only if a four-base gap was introduced into

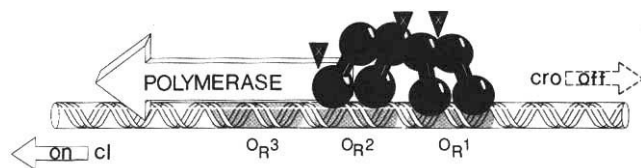


Fig. 1 Action of λ repressor. The λ repressor is both a negative and a positive regulator of gene expression. When bound to the right operator on the phage chromosome (O_R), the repressor shuts off (rightward) transcription of a gene required for lytic growth (*cro*) and, simultaneously, activates (leftward) transcription of its own gene (*cI*). Repressor bound to the strong site O_{R1} helps another repressor bind to the weaker adjacent site O_{R2} and this arrangement activates expression of the repressor gene as follows: the repressor at O_{R2} contacts polymerase and thereby helps it bind and begin transcription at the adjacent leftward promoter. At the same time rightward transcription is repressed because the bound repressors prevent polymerase from binding to the rightward promoter. Three protein-protein interactions are indicated by 'X' in the figure: each repressor dimer, the DNA-binding species, is held together primarily by contacts between carboxyl domains; interactions between carboxyl domains are also responsible for cooperative repressor binding to adjacent operator sites, but it is an amino domain of the repressor dimer at O_{R2} that contacts polymerase at the adjacent promoter. If the concentration of repressor increases transiently, site O_{R3} becomes occupied by a third repressor dimer and transcription of the repressor gene is turned off. (See ref. 7)

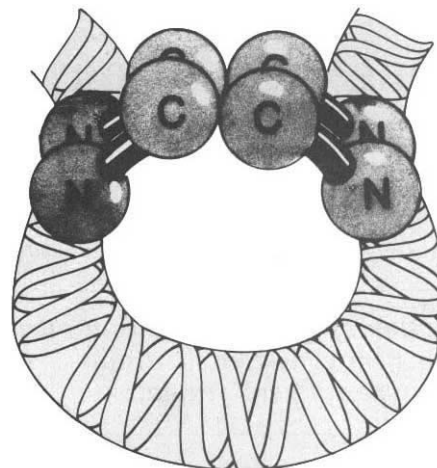


Fig. 2 λ repressors bound to two operator sites separated by seven helical turns. Although the precise points of contact between the dimers seem different from those indicated in Fig. 1, we suspect that either flexibility of the protein or some modification in the path of the DNA allows similar interactions in both cases⁸.

one of the DNA strands between the sites. This alteration presumably increases the torsional flexibility of the DNA and facilitates twisting, so that the proteins can be apposed. Whether the operator sites are adjacent or separated, cooperative binding of λ repressor depends upon the carboxyl domain of the protein, which does not contact DNA. These and other results suggest that λ repressors bind cooperatively to separated operator sites just as they do to adjacent sites, the DNA in between bending smoothly to accommodate the interaction as indicated in Fig. 2.

The idea that cooperative binding of λ repressor to separated sites induces a smooth bend in the DNA plausibly accounts for the following observation: as repressor binds, the DNA between the sites becomes alternately hypersensitive and resistant to DNase, the interval separating these recurring hypersensitive and resistant sites being five base pairs. We imagine that in the DNA between the bound repressors, the minor groove (the site recognized by DNase) is expanded along the outside surface of the loop and compressed along the inside surface (see Fig. 2), thereby periodically changing the DNase sensitivity. The pattern of DNase sensitivity resembles that observed with DNA wrapped around a protein⁹⁻¹¹ or in a nucleosome¹², but as confirmed by the electron microscopy experiments I describe below, the looped DNA in between the sites is not in contact with repressor. Perhaps the DNase-hypersensitive sites found near active genes in eukaryotic chromosomes¹³ are regions of distorted DNA structure caused by looping.

In collaboration with J. Griffith, we have now used electron microscopy to provide direct visual evidence for DNA looping induced by λ repressor binding¹⁴. The pictures show λ repressors bound to DNA fragments bearing two λ operators separated by 5 helical turns, measuring from the centres of the sites. The repressors are evidently in contact, the DNA in between looped out. The looped structures are not seen if the spacing between the operator sites is 4.5 or 5.5 helical turns. The results reinforce our conclusion that the bound repressors interact only if the operators are positioned so that the repressors are on the same side of the helix, a finding consistent with the assumption that the energy needed to twist DNA one-half turn (and thereby align repressors fixed to opposite sides of the helix) is prohibited.

Cooperative binding of λ repressors to separated sites, with concomitant DNA looping, requires no special DNA sequences

(as far as we know) and is readily observed on linear DNA under roughly physiological conditions using purified repressor and DNA. The repressor is not specially designed for looping because, as noted above, the operator sites are immediately adjacent on the ordinary λ chromosome. We do not yet know how increased separation between the sites (beyond seven turns) will affect the reaction, nor do we yet know whether supercoiling the DNA, or adding nonspecific DNA-binding proteins (for example, the HU protein of *E. coli*, or histones) will facilitate or hinder the reaction. Unpublished experiments (A. Hochschild and M.P., in preparation) show that the cooperative binding to separated sites observed *in vitro* can also be detected *in vivo* using appropriately designed DNA molecules.

New cases of gene regulation at a distance in bacteria are now coming to light^{15,16}. Even the Lac repressor, long thought to be a paradigm of single-site binding, can bind cooperatively to separated sites and there is some evidence that this binding may be physiologically important (J. Gralla, B. Muller-Hill, personal communications; see also ref. 17). Although there are cases involving the non-cooperative binding of regulatory proteins (for example the λ Cro protein) they may prove to be the exceptions. I have argued elsewhere that cooperative DNA binding of regulatory proteins is a useful strategy for increasing the specificity of DNA-protein interactions and for constructing sensitive genetic switches⁷. In some cases the protein binding sites are adjacent (as is ordinarily the case in λ and SV40) and in other cases they are separated, DNA looping allowing the same interactions as when the proteins are adjacent. Perhaps we have here another example of a common motif: viruses use their DNA more economically than do cells.

I now consider alternative models designed to explain action at a distance. Various experimental results, all consistent with looping, argue against the alternative models.

Twisting

One version of this model would have regulatory proteins bind directly to some altered form of DNA, for example left handed or single stranded; another version imagines that the regulatory protein has an enzymatic activity that alters DNA conformation, for example by unwinding it. Gene activation would then be a

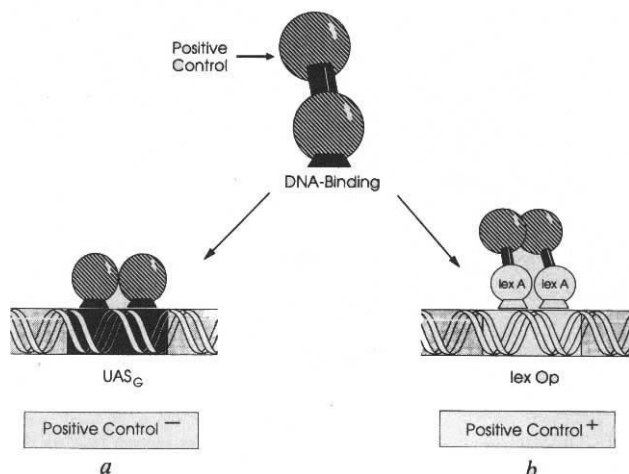


Fig. 3 Separation of the DNA-binding and positive control functions in a eukaryotic transcriptional regulator. The top line imagines the yeast activator GAL4 as consisting of two domains, the amino domain which recognizes sequences in the galactose upstream activating sequence (UAS_G) and the carboxyl domain which, when bound to DNA through the amino domain, contacts another protein to mediate positive control. In *a*, the carboxyl domain has been removed; this fragment recognizes UAS_G but fails to activate transcription. In *b*, the DNA-binding domain has been replaced with that of the bacterial protein LexA. This protein activates transcription in yeast if a *lexA* operator is positioned near the start site of a gene. The proteins are shown as binding to DNA as dimers but we are not certain this is correct^{25,26}.

consequence of conformation changes propagated through the DNA which allow other proteins to bind and begin transcription. The following considerations argue against these views.

A large number of known or suspected DNA-binding regulatory proteins have been isolated from both prokaryotes and eukaryotes and there is good reason to believe that all recognize ordinary helical DNA. For one class of these proteins, exemplified by λ repressor, the structure of the protein-DNA complex is known. Each of these proteins bears protruding α -helices that fit into the DNA major groove and make sequence-specific contacts with functional groups exposed along the edges of base pairs. As expected, these proteins do not greatly alter the structure of DNA upon binding (see ref. 18, for example). Other transcriptional regulators probably recognize specific DNA sequences using probes other than α -helices (see, for example, ref. 19) but there is no hint that these proteins recognize greatly altered forms of DNA. The known regulatory proteins recognize their sites in linear DNA with binding constants that approach those that describe the well-characterized Lac or λ repressor-operator interactions. Were these proteins to recognize some greatly altered DNA structure, this result would not have been obtained.

One specific suggestion, namely that enhancer sequences are recognized as a form of left-handed DNA called Z-DNA²⁰, now appears to be incorrect. Mutational analyses of enhancers fail to confirm the simple prediction that alternating purines and pyrimidines, the likeliest Z-forming sequences, are required for enhancer function²¹. The presumed correlation between active genes and Z-DNA in *Drosophila*, suggested by experiments with antibodies, has been demonstrated to be an artefact of sample preparation²².

If a transcriptional regulator does not change DNA structure, how does it activate transcription? For the case of λ repressor we know the answer: one surface of repressor contacts DNA and another contacts RNA polymerase and activates gene

expression. An important argument supporting this picture was the isolation of mutant λ repressors (called *pc* for positive control) that bind DNA normally but that fail to activate transcription²³. The changed amino acids in these mutants are found on that surface of repressor that had been predicted, on independent grounds, to approach most closely the adjacent RNA polymerase. The properties of these mutants thus show that although DNA binding of repressor is necessary for gene activation, it is not sufficient; an additional site, presumably a site of contact with polymerase, is also required.

We have extended these ideas to a eukaryotic case, showing that the DNA-binding and gene activation functions of a yeast transcriptional regulator can be separated. The protein we study, GAL4, activates transcription of the *GAL* genes in the presence of galactose. The protein binds to sequences in a control element called the UAS_G (the galactose upstream activating sequence) and activates transcription at sites about 250 base pairs away in wild-type strains and as far as 750 base pairs away in modified strains²⁴.

Figure 3 shows the properties of two fragments of GAL4, one of which (*a*) bears the GAL4 DNA-binding function and the other of which (*b*) bears the activation function. Protein fragment *a*, consisting solely of the first 98 amino acids of GAL4, binds the GAL4 recognition sites as assayed *in vitro* and *in vivo* but fails to activate transcription. A hybrid molecule bearing the amino-terminal portion of GAL4 fused to β -galactosidase has similar properties. These GAL4 derivatives, lacking amino-acid sequences from the carboxyl end of the molecule, are formally equivalent to the *pc* mutants of λ repressor²⁵. Protein *b*, bearing the carboxyl 800 amino acids of GAL4 attached to a DNA-binding domain of the bacterial repressor LexA, acts as a positive regulator in yeast in the absence of the ordinary GAL4 recognition sequences. The hybrid protein binds to *lexA* operators but not to UAS_G and its activity as a positive regulator requires that a *lexA* operator be positioned near the transcription start site of the gene²⁶.

These experiments argue that, for GAL4 as for λ repressor, the role of DNA binding is to position the protein on the DNA so that it can use some other activity to stimulate transcription. Although there is no indication that GAL4 itself can twist DNA, none of these experiments excludes the possibility that GAL4 helps some other protein to bind adjacent to UAS_G and that that protein activates transcription by untwisting (or twisting) DNA. However, the following experiment, performed with a mammalian virus, argues against this possibility.

Plon and Wang²⁷ have designed an incisive experiment that topologically separates the DNA site recognized by a regulatory protein (or proteins) from a gene without destroying gene activation. They constructed a 'tailed circle' (see Fig. 4) in which the regulatory sequence (the SV40 enhancer) forms a hairpin protruding from an otherwise intact circular DNA molecule that includes a gene (human β -globin). The strong result is that, when this construct is introduced into cells, transcription of the gene is enhancer dependent, as it is when gene and enhancer are disposed on ordinary DNA molecules. Thus, twisting of the enhancer cannot be responsible for gene activation, because twisting the protruding enhancer could have no effect on the topology of the gene.

Sliding

In this model a protein recognizes a specific site on DNA and then moves (slides, tracks) along the DNA to another specific sequence where, perhaps by interacting with another protein, it initiates transcription. Another version of essentially the same idea imagines that the regulatory protein remains bound at the original site and the DNA is threaded past or through the bound protein until the second critical site is encountered. (I am restricting the discussion to cases involving communication between two specific DNA sites. Where one specific and one nonspecific

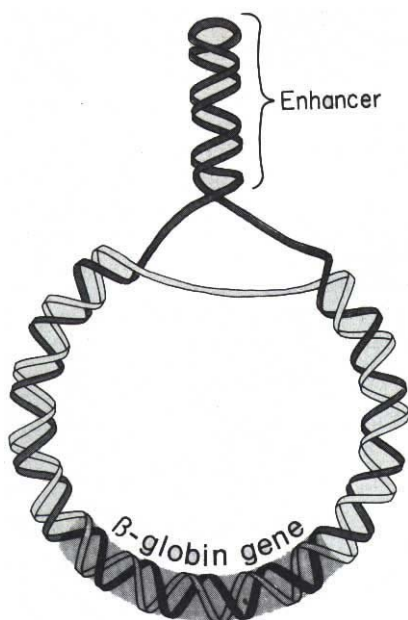


Fig. 4 The 'tailed circle' consists of an enhancer linked to, but topologically separated from, a gene. One of the DNA strands of this plasmid bears two copies of an SV40 enhancer sequence, one copy inverted with respect to the other. This extra region protrudes from the circle and self-pairs to form a functional enhancer. The main body of the circle contains the β -globin gene, transcription of which is increased by the enhancer. Twisting of the enhancer has no effect on the winding of the strands on the main body of the circle; nevertheless, the enhancer efficiently increases β -globin transcription²⁷.

site are involved, for example in the action of type I restriction enzymes, sliding may operate²⁸.)

No experiment involving transcriptional regulation, to my knowledge, directly argues against the sliders. But the strongest presumed example of *bona fide* sliding has been convincingly reinterpreted in the light of recent experiments. The cases in point involve two examples of site-specific recombination in bacteria. Both illustrate a remarkable common property of various site-specific recombination systems: two sites separated by many thousands of base pairs recombine only if the two sites are in a specified orientation; if one of the sites is inverted with respect to the other no recombination is observed. How does a protein bound at one site 'sense' the orientation of another which is 50,000 base pairs away? Tracking seemed a reasonable explanation and this mechanism could explain other experimental findings²⁹.

Craigie and Mizuuchi³⁰ have now eliminated tracking as a mechanism for site-specific recombination involving the ends of the bacteriophage Mu. This recombination reaction proceeds *in vitro* at high efficiency with purified proteins and ATP. Craigie and Mizuuchi³⁰ show that two Mu sites not on the same DNA molecule can nevertheless recombine if they are part of separate molecules that are intertwined (Fig. 5); in this case a protein could not track from one site to the other. These findings are consistent with tests of the tracking model involving the bacterial transposon Tn3, in which Benjamin *et al.*³¹ used electron microscopy, as well as other techniques, to analyse products of recombination *in vitro*. They found a distribution of structures inconsistent with the tracking model.

Boocock *et al.*³² have also described experiments which strongly support the idea that the ends of Tn3 find themselves by looping, and they have described a plausible model, which I shall not review here, that explains the orientation dependence

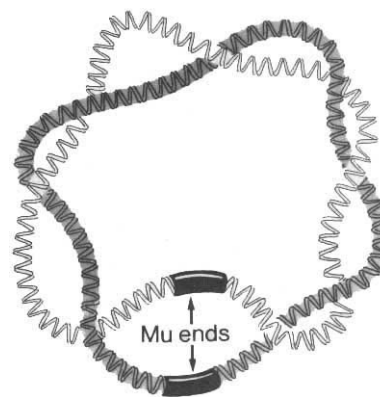


Fig. 5 Two intertwined DNA circles show that recombination is possible without tracking. Each line represents a double-stranded DNA molecule and each bears a Mu phage end. These sites efficiently recombine *in vitro* under the direction of site-specific recombination proteins. Because the sites are on separate molecules, it is impossible to travel from one to the other by tracking³⁰.

of the reaction. These findings help to unify our understanding of site-specific recombination; in another case, that involving the ends of the λ prophage, it is already well accepted that the sites find each other by looping^{33,34}.

Oozing

The simplest form of this model is that the binding of a regulatory protein to its operator helps binding of another protein to adjacent sequences, which in turn helps another to bind next to it, and so on, until a procession of proteins has oozed out from the control sequence to the gene where transcription is initiated. This idea as applied to transcription cannot be discounted entirely, and indeed some oozing around the control sequence or around the transcription start site seems a possibility (see, for example, the line-up of proteins in Fig. 1). But it seems implausible to imagine such an effect working over thousands of base pairs. The idea under consideration here should not be confused with nonspecific wrapping of DNA around histone-like proteins.

Oozing is not important for site-specific recombination. Experiments performed *in vitro* show that for the case involving the ends of Tn3, the amount of protein sufficient to catalyse the reaction is insufficient to cover the DNA between the sites³⁵.

Generalizations

If the thrust of these arguments is correct, we have the outlines of a unified view of gene regulation. In its simplest form this view would make two statements. First, regulatory proteins recognize specific sequences in DNA using structures that are complementary to the ordinary helix. The α -helix that protrudes from the surface of λ repressor and explores functional groups exposed in the major groove of DNA is an example of such a structure. Second, DNA-bound regulatory proteins influence transcription by excluding binding of other proteins (for example by competing with polymerase for an essential site) or, more generally, by touching another DNA-bound protein. In the λ case, for example, one DNA-bound repressor helps another to bind, and the latter in turn helps polymerase to bind and begin transcription, and these cooperative effects are mediated by protein-protein interactions. The point I have explicitly argued in this review is that these descriptions hold whether the interacting proteins are adjacent on DNA or whether they are separated—in the latter case the DNA in between the separated sites loops out to allow the protein-protein interactions.

Should looping prove to be a general phenomenon in gene regulation at a distance, it is not difficult to imagine how various forms of positive and negative control might be effected. For example, a protein bound to a distal site might help another protein bind to a proximal site and the latter might then directly activate or repress a gene. Alternatively, a bound activator protein might be constrained and rendered inactive by interaction with another protein bound at a distal site. Proteins binding within the loop might break it, thereby relieving the negative or positive effect of the interaction. Proteins bound to the same

DNA sites might have quite different effects on gene regulation depending upon the nature of their interaction with other proteins. Our problem is no longer to invent a possible mechanism; rather, it is to see how far one operative idea, looping, will carry us.

I thank students and colleagues at Harvard; also Ben Lewin, Howard Nash and Bob Schleif for helpful comments and Carol Morita for the figures. This work has been supported in part by NIH grant GM22526 to M.P.

1. Irani, M. H. Orosz, L. & Adhya, S. *Cell* **32**, 783-788 (1983).
2. Majumdar, A. & Sankar, A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6100-6104 (1984).
3. Dunn, T., Hahn, S., Ogden, S. & Schleif, R. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5017-5020 (1984).
4. Hahn, S., Hendrickson, W. & Schleif, R. *J. molec. Biol.* **188**, 355-367 (1986).
5. Martin, K., Huo, L. & Schleif, R. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3654-3658 (1986).
6. Takahashi, K. *et al. Nature* **319**, 121-126 (1985).
7. Ptashne, M. *A Genetic Switch* (Cell & Blackwell Scientific Press, Cambridge & Palo Alto, 1986).
8. Hochschild, A. H. & Ptashne, M. *Cell* **44**, 681-687 (1986).
9. Kirkegaard, K. & Wang, J. C. *Cell* **72**, 721-729 (1981).
10. Morrison, A. & Cozzarelli, N. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1416-1420 (1981).
11. Sawadogo, M. & Roeder, R. *Cell* **43**, 165-175 (1985).
12. Noll, M. *J. molec. Biol.* **116**, 49-71, (1977).
13. Wu, C., Bingham, P. M., Livak, K. J., Holmgren, R. & Elgin, S. C. R. *Cell* **16**, 797-806 (1979).
14. Griffith, J., Hochschild, A. & Ptashne, M. *Nature* **322**, 750-752 (1986).
15. Dandanell, G. & Herman, K. *EMBO J.* **4**, 3333-3338 (1985).
16. Reitzer, L. J. & Magasanik, B. *Cell* **45**, 785-792 (1986).
17. Mossing, M. & Record, T. *Science* (in the press).
18. Anderson, J., Ptashne, M. & Harrison, S. *Nature* **316**, 596-601 (1985).
19. Miller, J., McLachlan, A. D. & Klug, A. *EMBO J.* **4**, 1609 (1985).
20. Nordheim, A. & Rich, A. *Nature* **303**, 674-679 (1983).
21. Zenke, M. *et al. EMBO J.* **5**, 387-397 (1986).
22. Stollar, D. *CRC crit. Rev. Biochem.* **20**, 1-36 (1986).
23. Hochschild, A., Irwin, N. & Ptashne, M. *Cell* **32**, 319-325 (1983).
24. Giniger, E., Varnum, S. M. & Ptashne, M. *Cell* **40**, 767-774 (1980).
25. Keegan, L., Gill, G. & Ptashne, M. *Science* **231**, 699-704 (1986).
26. Brent, R. & Ptashne, M. *Cell* **43**, 729-736 (1985).
27. Plon, S. & Wang, J. *Cell* **45**, 575-580 (1986).
28. Yuan, R. A. *Rev. Biochem.* **50**, 285-315 (1981).
29. Krasnow, M. A. & Cozzarelli, N. R. *Cell* **32**, 1313-1324 (1983).
30. Craigie, R. & Mizuuchi, K. *Cell* **45**, 793-800 (1986).
31. Benjamin, H. W., Matzuk, M. M., Krasnow, M. A. & Cozzarelli, N. R. *Cell* **40**, 147-158 (1985).
32. Boocock, M. R., Brown, J. L., Sherratt, D. J. *Biochem. Soc. Trans.* **14**, 214-216 (1986).
33. Mizuuchi, K., Gellert, M., Weisberg, R. A. & Nash, H. J. *J. molec. Biol.* **141**, 485-494 (1980).
34. Spengler, S. J., Stasiak, A. & Cozzarelli, N. R. *Cell* **42**, 325-334 (1985).
35. Krasnow, M. A., Matzuk, M. M., Dungan, J. M., Benjamin, H. W. & Cozzarelli, N. R. in *Mechanisms of DNA Replication & Recombination* (ed. Cozzarelli, N. R.) 637-659 (Liss, New York, 1983).

ARTICLES

Melting history of Antarctica during the past 60,000 years

L. D. Labeyrie*, J. J. Pichon[†], M. Labracherie[†], P. Ippolito⁺,
J. Duprat[†] & J. C. Duplessy*

* Centre des Faibles Radioactivités, Laboratoire Mixte CNRS/CEA, Domaine du CNRS, 91190 Gif-sur-Yvette, France

[†] Département de Géologie et Océanographie, L. A. 197, Université de Bordeaux 1, Avenue des Facultés, 33405 Talence, France

Marked changes in the surface-water hydrology of the Southern Ocean during the past 60 kyr are revealed by a detailed comparison of the oxygen isotopic composition of planktonic and benthic foraminifera from sediment cores and the surface-water temperature estimated by a transfer function derived from the distribution of diatoms in the same sediments. From 35 to 17 kyr BP, the Southern Ocean polar front was covered by a melt-water lid containing a significant contribution from melting icebergs, calved from Antarctic ice shelves. These icebergs may have originated from a succession of surges of the ice shelves.

THE melting history of Antarctica during the most recent climatic cycle is a challenging problem which has an important bearing on global climatic models¹, on the formation of bottom and intermediate water, and consequently on the CO₂ content of the atmosphere². One open question is the possibility of large surges of part of the Western Antarctic ice sheet³. Such abrupt events should be detectable using the oxygen isotopic composition of fossil planktonic foraminifera or other epipelagic organisms, because Antarctic ice, due to its conditions of accumulation, is depleted by 35-50‰ in δ¹⁸O compared to mean ocean water. However, as the foraminiferal isotopic ratio is a function of both the temperature and the isotopic composition of the surrounding water, we must estimate the water temperature independently. Imbrie and Kipp⁴ have shown that surface-water temperature may be accurately estimated by transfer functions based on quantitative micropalaeontological analyses. Using a transfer function derived from Southern Ocean diatoms⁵, we may estimate independently the two components of the oxygen

isotopic record of the epipelagic foraminifera. The oxygen isotopic ratio of the benthic foraminifera analysed at the same levels in sediment cores gives the fraction of the sea-water isotopic changes which affects the whole water column, this being connected with the evolution of the continental ice sheets. We may thus estimate the local surface-water isotopic changes. We present here the results of such a study, carried out on two sediment cores from the southern Indian Ocean which cover the period including isotopic stages 3 and 2 of the last glacial (ref. 6), and the following deglaciation.

¹⁸O results and chronology

The cores MD73026 (44°59' S, 53°17' E) and MD84527 (43°49' S, 51°19' E) are located in the high-productivity belt north of the Crozet Plateau. Their position within the polar front system⁷, the major boundary between subtropical and Antarctic surface waters, makes them ideal for a study of hydrographic changes in the Southern Ocean. The high sedimentation rate

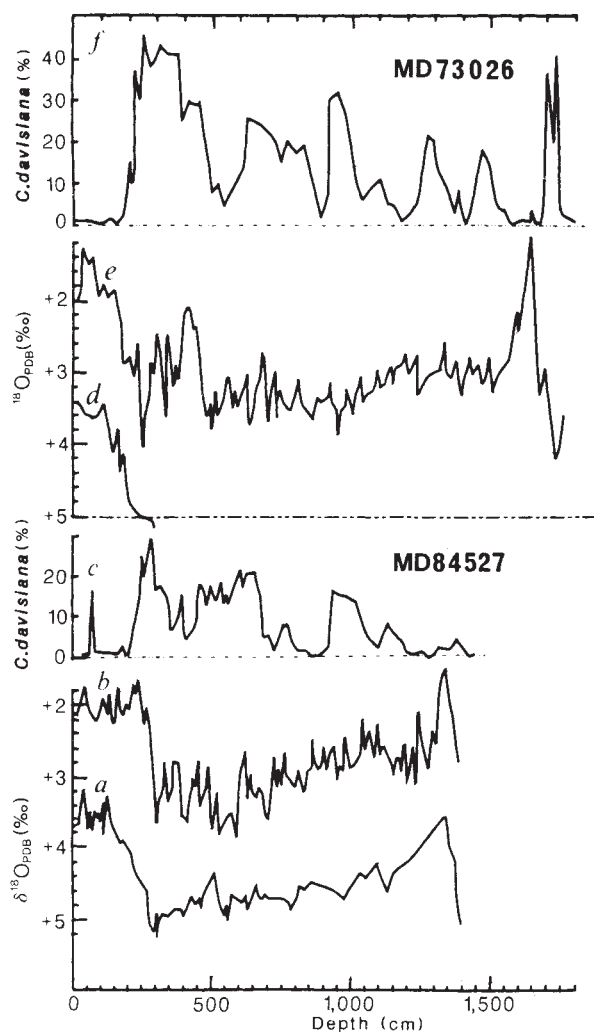


Fig. 1 *a–c*, Records from core MD84527 (43°49.3' S, 51°19.1' E; 3,262 m depth). *a*, $\delta^{18}\text{O}$ (in parts per 10^3 , relative to the PDB standard) ($\delta^{18}\text{O} = [(R_{\text{carbonate}}/R_{\text{standard}}) - 1] \times 10^3$, where $R = \{^{18}\text{O}/^{16}\text{O}\}$) of the benthic foraminifera *Cibicides wuellerstorfi*, plotted against depth in the core (corrected by +0.64‰ to take into account the specific fractionation⁴⁸). *b*, $\delta^{18}\text{O}_{\text{PDB}}$ for the planktonic foraminifera *Neogloboquadrina pachyderma* l.c. *c*, In the same core, relative amount of the radiolarian species *Cycladophora davisiana*. *d–f*, Records from core MD73026 (44°59' S, 53°17' E; 3,429 m depth). *d*, $\delta^{18}\text{O}_{\text{PDB}}$ of the benthic foraminifera *Nonion* sp., corrected by +0.4‰ for specific fractionation. *e*, $\delta^{18}\text{O}_{\text{PDB}}$ for *N. pachyderma* l.c. *f*, In the same core, relative amount of *C. davisiana*.

(> 8 cm kyr⁻¹) has provided a highly detailed record of the most recent climatic cycle. We have analysed in these cores the oxygen isotopic ratio of the epipelagic foraminiferal species *Neogloboquadrina pachyderma* left-coiled and, wherever possible, the benthic species *Melonis barleanum* and *Cibicides wuellerstorfi*. We have also analysed the foraminiferal $\delta^{13}\text{C}$ and the diatom $\delta^{18}\text{O}$ values, the diatom, foraminiferal and radiolarian specific abundances, and the abundance of ice-rafted detrital grains (L.D.L. *et al.*, in preparation). Figure 1 shows the relative abundance of the radiolarian species *Cycladophora davisiana*, a good stratigraphic marker in the Southern Ocean^{8,9}, together with the foraminiferal $\delta^{18}\text{O}$ values, as a function of depth in the cores. These records are similar to the published records of nearby cores MD73025^{10–12} and RC11–120⁸, if we take into account the varying sedimentation rates.

We have referred the records of MD84527 and MD73026 to the SPECMAP time scale¹³, based on RC11–120, to allow a

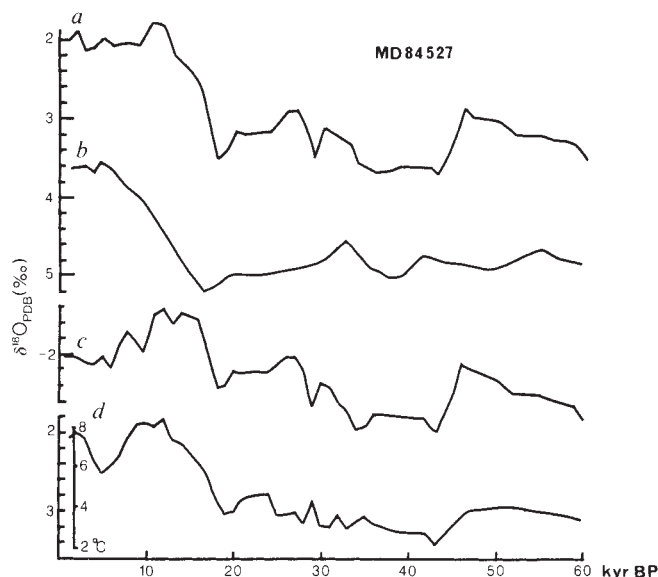


Fig. 2 Records from core MD84527, plotted against time. *a*, Planktonic isotopic record; *b*, benthic isotopic record; *c*, local isotopic changes ($\delta^{18}\text{O}$ (planktonic) – $\delta^{18}\text{O}$ (benthic)), expressed on the same timescale. A correction factor of +0.4‰ has been added to all the values in the period 0–9 kyr BP to correct for the effect of bottom-water temperature change (see text). *d*, Surface-water temperature change at the location of the core. To allow direct comparison with the other records, the isotopic effect of these temperature changes is plotted on the $\delta^{18}\text{O}$ scale of *N. pachyderma*, using the calibration of Table 1.

better comparison between them and with other records. This has been done by adjusting, at each depth level in the three cores, the foraminiferal $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ values, and the relative abundance of *C. davisiana*. The accuracy of the correlation is better than 1 kyr during deglaciation, but only 3–5 kyr for isotopic stages 2 and 3. This is not critical for the present study.

The benthic foraminiferal record of core MD73026 covers only the period from 0 to 18 kyr BP. However, whereas the North Atlantic Ocean experienced significant deep-water temperature changes during the most recent glacial period¹⁴, the isotopic records of benthic foraminifera during the most recent climatic cycle are congruent in the Indian and Pacific Oceans¹¹. We have therefore used the $\delta^{18}\text{O}$ record of *C. wuellerstorfi* in core MD84527 as a reference for core MD73026. The resulting foraminiferal $\delta^{18}\text{O}$ records of cores MD84527 and MD73026 are plotted against age in Figs 2 and 3.

The surface-water isotopic signal

Significant differences in the isotopic records of the planktonic and benthic foraminifera are evident in both cores. Note in particular the earlier (by 2–3 kyr) decrease in the isotopic ratio of the planktonic foraminifera during deglaciation (17–10 kyr BP), and the anomalously low isotopic ratio of these foraminifera before and during the glacial maximum (35–17 kyr BP). Such differences also exist in the nearby core MD73025¹⁵. The sediments in these three cores are not obviously disturbed, and they contain normal records of the other stratigraphical indexes (benthic foraminifera $\delta^{18}\text{O}$, radiolarian species *C. davisiana* and diatom species *Eucampia antarctica*). The effects of bioturbation on the isotopic records have been recently quantified by ¹⁴C dating of the same monospecific samples used for isotopic analyses¹⁶; we conclude that bioturbation cannot explain the observed isotopic anomalies.

The differences in the isotopic records between surface (planktonic) and deep (benthic) foraminifera must be explained by

Table 1 Data used for calculations of oxygen isotopic fractionation between *N. pachyderma* and water

Sample	Lat.	Long.	S (‰)	$\delta^{18}\text{O}_{\text{water}}$	T (°C)	$\delta^{18}\text{O}_{\text{pachy}}$
MD73026	44°59' S	53°17' E	34.40	+0.3	+8.5	2.00
V2786	66°36' N	01°07' E	35.15	+0.55	+6	2.60
V2838	69°23' N	04°24' W	35.1	+0.5	+6	2.93
K11	71°47' N	01°36' E	35.1	+0.5	+5	2.94
V2760	72°11' N	08°35' W	35.05	+0.45	+5	2.84
HU7541	62°39' N	53°52' W	34.5	+0.1	+5	3.00
HU7542	62°39' N	53°54' W	34.5	+0.1	+5	2.58
MD80304	51°04' S	67°44' E	33.80	-0.15	+3.5	2.91
MD84551	55°00' S	73°16' E	33.90	-0.15	+1.5	3.15
CH7707	66°36' N	10°30' W	34.9	+0.4	+1	3.68
MD82424	54°05' S	00°21' W	33.93	-0.15	+1	3.59
M269-965	60°54' S	57°06' W	33.60	-0.15	-0.25	3.9

$\delta^{18}\text{O}_{\text{PDB}}$ of the planktonic foraminifera *N. pachyderma* *sen.* in samples extracted from core tops, and a single sediment trap sample (M269-965)⁴⁵. The temperatures (T) and salinities (S) are the mean summer values extracted from atlases^{46,47} for 80–100 m water depth at the sample locations. The water $\delta^{18}\text{O}$ values are calculated from the salinities²³.

differences in the evolution of the surface and deep waters in the Southern Ocean during the past 60 kyr. Both the temperature and the isotopic composition of the waters may be involved, and these must be estimated independently. Duplessy *et al.*¹⁰ have used the benthic record of MD73025 as a reference for global sea-water isotopic changes during the accumulation and melting of the continental ice sheets. More recent results¹⁴ indicate that of the 1.6‰ total isotopic difference between the last glacial maximum and the Holocene in that core, 0.4‰ has to be attributed to a 1.5 °C warming of the deep Indian Ocean water at ~9 kyr BP, when the present mode of formation of the North Atlantic Deep Water developed. Accordingly, we adopt the benthic foraminiferal record of core MD84527 for the past 60 kyr: as a reference for the isotopic record for mean ocean water, after correcting by +0.4‰ all isotopic values from the period 0–9 kyr BP. By subtracting the corrected benthic signal from the *N. pachyderma* records of MD84527 and MD73026, we obtain the 'local' isotopic records for the planktonic foraminifera (Figs 2c, 3c).

These records integrate the effect of the surface-water temperature change on the isotopic fractionation during carbonate precipitation. The first step in estimating this effect is to reconstruct the temperature changes during the past 60 kyr at the location of the cores. This has been done using a transfer function derived by multivariate correlation⁴ between the diatom specific abundance in Holocene core tops from the Southern Ocean and the present surface-water temperature at the location of the cores⁵. The summer (February) surface temperature is calculated by this method with an accuracy of ± 1.1 °C.

The temperature records obtained by this method (Figs 2d, 3d) show two interesting features. First, the increase in surface-water temperature preceded by ~2–3 kyr the changes in isotopic ratio of the benthic foraminifera, which marks the initiation of the Northern Hemisphere deglaciation. This confirms the results of Hays *et al.*¹⁷, who have shown that the temperature maximum in the Southern Ocean core RC11-120 was near 9.4 kyr BP, that is, leading by ~2 kyr the Northern Hemisphere Holocene optimum. Secondly, the summer temperature was fairly stable in core MD84527 (between +2 and +4.5 °C) throughout the period 17–60 kyr BP. This means that the Antarctic polar front remained at the latitude of this core throughout this period. The temperature record of MD73026 is presently available only until 35 kyr BP, thus limiting the quantitative interpretation of the isotopic anomaly in that core.

The temperature dependence of the oxygen isotopic fractionation during carbonate precipitation in *N. pachyderma* was obtained by comparing *N. pachyderma* isotopic analyses from 12 Holocene core tops with the present summer water temperature at the location of these cores (Table 1). The depth

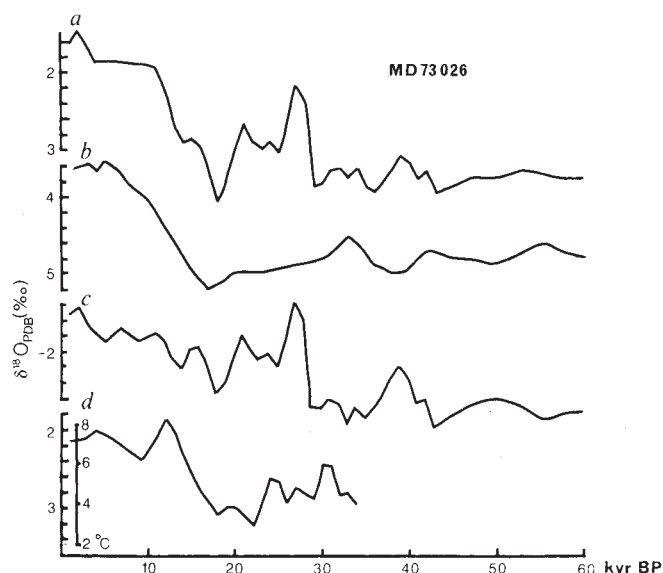


Fig. 3 $\delta^{18}\text{O}_{\text{PDB}}$ of the foraminifera in core MD73026, plotted against time. *a*, Planktonic foraminifera *N. pachyderma* l.c.; *b*, benthic record of core MD84527; *c*, local isotopic changes for core MD73026 (as for Fig. 2c); *d*, summer surface-water temperature changes at the location of the core (as for Fig. 2d).

habitat of this species in low-productivity areas is in the 80–100 m range¹⁸, of the same order or slightly deeper than that of the diatoms¹⁹. In the higher-productivity area of the North Pacific, *N. pachyderma* l.c. occupy the entire surface mixed layer²⁰. Regression analysis over the temperature range -1 to +9 °C yields $\delta^{18}\text{O}_{\text{pachy}} - \delta^{18}\text{O}_{\text{water}} = 3.84 - 0.25T$ with 1σ errors of 0.12 in the constant and 0.03 in the slope. These results are in agreement with the estimate of Shackleton²¹ for benthic foraminifera living in this temperature range. Using the relationship obtained for *N. pachyderma*, we may now calculate the isotopic changes which would have been due solely to the changes in water temperature. The corresponding $\delta^{18}\text{O}$ scales are reported beside the temperature scales in Figs 2d and 3d.

The variations of the surface-water $\delta^{18}\text{O}$ have been calculated by correcting the local isotopic records of the planktonic foraminifera (Figs 2c, 3c) for the effect of temperature changes. These estimates of the net surface-water isotopic changes are shown in Fig. 4a, b. Although there is some variability, the general trend in the signal corresponds to a negative shift in the surface-water isotopic ratio for the entire glacial period compared to the present. Several shifts of large amplitude (1‰) occur between 17 and 30–35 kyr BP; these are particularly evident in core MD73026 (Fig. 4a).

Low- $\delta^{18}\text{O}$ water from melting icebergs

Several mechanisms may be invoked to explain the decrease in the surface-water isotopic ratio. There is today a 0.4‰ negative gradient in the surface-water isotopic ratio from north to south through the Antarctic polar front^{23,24}. The subtropical water is enriched in ^{18}O by evaporation (+0.2‰ relative to standard mean ocean water, SMOW), and surface Antarctic water is depleted by excess precipitation (-0.2‰). Studies in progress using global climatic models with isotopic tracers indicate that neither this isotopic gradient, nor the precipitation over the Southern Ocean, could have increased significantly during the most recent glacial (J. Jouzel, in preparation). We have therefore to look for a source of water with a much lower isotopic ratio than that of the rain and snow which today fall over the Southern Ocean. Melting sea ice is of no effect, because there is very little isotopic fractionation during freezing of sea water²⁵. On the other hand, melting icebergs could be of significant influence; their isotopic ratio, depending on the source of ice within Antarctica, lies

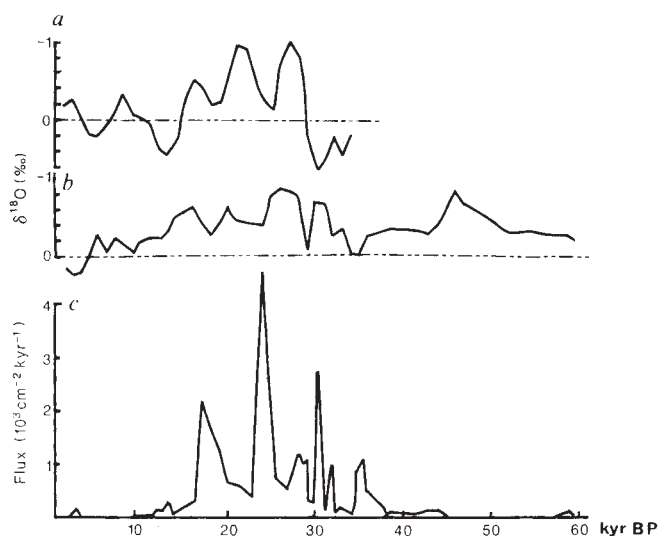


Fig. 4 *a, b*, Net surface-water isotopic variations for cores MD73026 (*a*) and MD84527 (*b*), obtained by subtraction of the surface-temperature signal from the local isotopic changes. *c*, Absolute fluxes of glacial detrital quartz grains in core MD84527.

between -35 and -55% (ref. 25). To date, however, such an input of iceberg melt water has been measured only along the Filchner Ice Shelf, in the Weddell Sea²⁵. To explain the isotopic anomalies, we must therefore consider periods of major increase in the production and melting of icebergs during the last glacial.

The first piece of evidence in support of this hypothesis is the distribution of ice-rafted detritus (IRD). The relationship between glaciation and the abundance of IRD in the sediments of the Southern Ocean is well known^{9,26-28}. Keany *et al.*²⁸ have connected the north-south shifts in IRD deposition to the north-south shifts in the iceberg melting zone. Although a large part of the IRD deposited between the Scotia Sea and Kerguelen Islands consists of tholeiitic glass shards transported by sea ice and icebergs from the South Sandwich Islands^{26,27}, quartz and chlorite probably originate only from the old continental crust of Antarctica. Their incorporation in icebergs could have been facilitated by the extension of the grounded ice shelves over the continental margins. To test for a possible increase in the amount of IRD transport by Antarctic icebergs, we have estimated the fluxes of quartz grains during the last glacial in core MD84527. We sub-sampled known sediment volumes, and counted all the quartz grains in the size fraction larger than $63\ \mu\text{m}$, thereby excluding wind-transported detritals. As shown in Fig. 4c, the quartz flux increases dramatically during the period from 17 to 35 kyr BP, that is, at the time of the major isotopic anomaly in both cores. Small differences in the relative depths of the IRD maxima and isotopic anomalies may be explained by several factors. The sampling for isotopic and IRD studies was done on separate occasions, at different levels in the core. Furthermore, the wet density of the IRD is much higher than for foraminifera and diatoms, thus allowing for differential settling of a few centimetres within the water-rich upper sediment before compaction. The earlier isotopic anomaly (at ~ 50 kyr BP) recorded in core MD84527 but not in the planktonic record of core MD73026 (Figs 2a, 3a) does not correspond to any significant increase in detrital quartz.

Other observations may help to explain the surface hydrography during these periods. Cooke and Hays⁹ have remarked that during the last glacial maximum (broadly defined as the *C. davisiana* zone b), IRD fluxes were large in both the Atlantic and south-west Indian sectors of the Southern Ocean. Before this time (*C. davisiana* zones c-d, ~ 40 – 60 kyr BP), the area of large IRD fluxes was limited to the Atlantic sector, a result

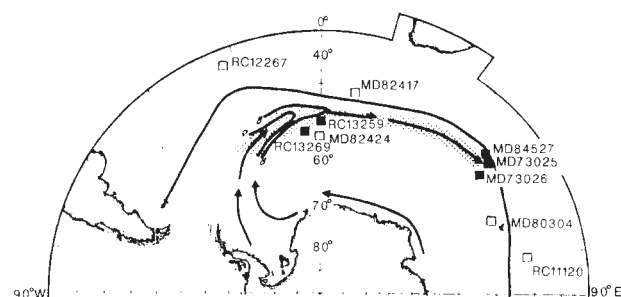


Fig. 5 Distribution of the surface-water isotopic anomaly at 18–20 kyr BP. ■, Cores containing the anomaly in the planktonic foraminifera or diatom records²²; □, cores with normal isotopic records; →, estimated trajectory of the icebergs extrapolated from present trajectories⁴⁹; —, position of the polar front at 18 kyr BP⁸. The shading indicates regions of high loadings at 18 kyr BP for factor 2 of Burckle's diatom factorial analysis, which is mostly composed of *E. antarctica*, and indicates surface-water low salinity from melting sea-ice and icebergs.

which agrees with our quartz flux study. However, for both these periods, and in the different sectors, there was a significant increase in the relative amount of *C. davisiana*. In its present habitat in the Okhotsk Sea, this species is associated with specific hydrographic conditions, namely a surface water lid of less saline water with a strong temperature minimum near its base, and the formation in winter of sea ice which melts in the summer²⁹. Morley and Hays²⁹ have proposed this hydrological situation as a modern analogue for the periods of high abundance of this radiolarian, and specifically for the last glacial in the Southern Ocean. The diatom *Eucampia antarctica* is abundant today around the Antarctic peninsula. Burckle³⁰ associates this species with neritic and/or low-salinity waters (melting sea ice and icebergs). The surface-water conditions favourable for its development are thus the same as those for *C. davisiana*. Burckle³⁰ interprets the distribution of *E. antarctica* in the sediment of the last glacial maximum around Antarctica as indicating an extension of a low-salinity belt south of the polar front. In our set of cores, there is also a good correlation between the relative abundance of *C. davisiana* (Fig. 1), several species of diatoms (*E. antarctica* and *Thalassiosira antarctica*) and the isotopic anomalies²². Figure 5 shows the distribution of the *E. antarctica* factor of Burckle³⁰, together with the cores with and without the isotopic anomaly, at ~ 18 kyr BP. This distribution corresponds to the high IRD fluxes at that period⁹. We interpret this data to be an indication of the trajectory of melting sea-ice and icebergs from the Weddell Sea south of the polar front, along the circumpolar drift. With this hypothesis, the lack of detrital quartz and other IRD in the south-west Indian sector during the surface-water isotopic anomalies of the 40–60 kyr BP period would indicate that most of the icebergs melted faster at that time, within the Atlantic sector. Cooke and Hays⁹ consider that sea ice was more restricted at ~ 50 kyr than at 18 kyr BP, thus allowing faster melting. We do not have data to substantiate this argument, and therefore will not discuss this early part of the last glacial.

The magnitude of the ice flux from Antarctica necessary to give the 1‰ isotopic anomaly may be estimated from Fig. 5. Assuming that the ice isotopic ratio was -40% , the water mixed-layer thickness during periods of melt-water input was 60 m, and the surface covered by the melt water was $2 \times 10^6\ \text{km}^2$ (30°W to 60°E , by 4° latitude), we need $10,000\ \text{km}^3$ of melt water from Antarctic ice. This number should be compared with the present-day volume of $2,000\ \text{km}^3$ melt water produced annually from melting Antarctic ice (refs 31, 32). It is not possible, therefore, to attribute the isotopic anomalies recorded during the last glacial to a regular calving, at the present rate, of icebergs from the ice shelves.

A much larger accumulation rate of the Antarctic ice sheet during the last glacial is most unlikely. The ice flow models developed to explain the data from the Dome C³³ and Vostok³⁴ ice cores imply a smaller ice accumulation rate during the last glacial. This is confirmed by the higher ¹⁰Be specific activity observed in the Vostok ice core at that time³⁵. The higher gas content of the Bird ice core during the last glacial maximum also indicates a lower altitude of ice formation in the central part of the West Antarctic ice sheet³⁶. There are conflicting land-based geological observations suggesting that the ice sheet could have had a larger extension and was also thicker at the periphery during the last glacial³⁷. It is thus possible that during at least part of the last glacial, accumulation rates were significantly larger than today around the periphery of Antarctica, but not in the more central parts. Taking into account the estimated volume of the total Antarctic ice sheet during the glacial period of $3.67 \times 10^7 \text{ km}^3$ (ref. 38), an erosion rate of $\sim 10^6 \text{ km}^3 \text{ yr}^{-1}$ could not be sustained for several kyr without being recorded in the Dome C and Vostok cores. The constraints are not so strong for the West Antarctic ice sheet. The lower altitude of ice accumulation during the glacial³⁶ could be interpreted as being due to a greater erosion rate than at present. Additional contributions from icebergs from the coastal ice sheets of the Antarctic Peninsula and the South Shetland and South Orkney Islands³⁹ are probable, but not quantifiable. Moreover, due to their proximity to the evaporation zones, these small ice caps probably had oxygen isotopic ratios similar to that of normal precipitation and therefore had a minor effect on the observed isotopic anomaly.

These arguments lead us to conclude that erosion of the Antarctic ice sheet could not have produced the volume of icebergs necessary to induce the observed isotopic anomalies as a more-or-less permanent regime during the last glacial. However, the detailed record of core MD73026 (Fig. 4a) does indicate that the ice output was irregular, with three well-defined peaks, one approximately every 6 kyr. We do not have a sufficiently precise chronology to estimate the duration of each melt-water event. The chronostratigraphy used in Fig. 3 is defined by a small number of references to the SPECMAP timescale (18, 21, 29, 35, 43 and 57 kyr BP, for isotopic stages 2 and 3), with the sedimentation rate taken as constant within each interval. However, Cooke and co-workers^{26,40} have demonstrated that the peaks of high IRD coincide with periods of high productivity of several species of radiolarians, and, indeed, our cores also have high accumulation rates of radiolarians which correlate with the high IRD and the melt-water events. Icebergs, due to their large size both above and below sea level, create small 'island effects' while drifting under the mixed influence of wind and currents. Reaching the area of high winds and rough seas (45–50° S), they will contribute to the vertical mixing, and hence increase the primary productivity. Neshyba⁴¹ has pointed out the importance of the iceberg-induced upwellings in the Weddell Sea. On the other hand, productivity is much smaller under a semi-permanent sea-ice cover⁸. It is therefore probable that sedimentation rate (both detrital and biogenic) was significantly higher during the melt-water events than it was between them. A duration of 1 kyr for each melt-water event, separated by 5 kyr of re-equilibration of the ice profile, is compatible both with the observed isotopic effects and with the present rate of ice accumulation. If similar phenomena were occurring at other places around Antarctica (especially east of the Ross Sea), each of the events would have had to be proportionately shorter to keep within the hypothetical constraint of a constant mean ice accumulation rate of $2,000 \text{ km}^3 \text{ yr}^{-1}$.

We therefore propose that the isotopic anomalies we observe record a succession of pulse-like surge events. Denton and Hughes³⁷ have stressed the importance for marine-based ice sheets of the removal of material by ice streams ('drawdown'). This mechanism feeds most of the present ice shelves in Antarctica. A large extension of the peripheral floating ice shelf around

Antarctica during the last glacial could have dramatically increased the volume of ice eroded by drawdown around the margins of the Antarctic ice sheet. Such a system would be semi-oscillatory, as ice streams would remove ice much more quickly than it could be replenished from the main body of the ice sheet. Within a few hundred years, all of the ice available at the head of the ice shelf would be removed, thus draining the coastal glaciers and ice streams. They will be reactivated only after a sufficient part of the ice sheet becomes once again marine-based. Similar models, with changes of smaller amplitude, have been discussed with respect to the present evolution of western Antarctica^{42,43}.

Denton and Hughes³⁷ have also stressed the teleconnections between Northern and Southern Hemisphere marine-based ice sheets. Their floating periphery is very sensitive to destabilization by sea-level changes. Moreover, each of the isotopic anomalies would represent an increase of several metres in global sea level, at the time when the boreal ice sheet was near its maximum extension. This may have helped to destabilize the marine-based ice sheets of the Northern Hemisphere, thus providing the amount of melting necessary to explain the high sea-level stands known around 25–30 kyr BP⁴⁴. A precise mapping of the net surface-water isotopic composition over the different oceans during the 17–30 kyr BP period, using a methodology similar to that presented here, should help to solve that problem.

Conclusions

Our results give new insight into the dynamics of the Antarctic ice sheet during the last glacial period. It is generally accepted that the Antarctic ice sheet had characteristics similar to those seen today³⁸, with a lower accumulation rate³⁴ and a surface temperature which responded in phase with global sea-level changes^{33,34}. We have demonstrated that the interactions between sea level, sea surface temperature around Antarctica, and the erosion rate of the Antarctic ice sheet may be more complicated. There is a definite indication that the Southern Ocean led by 2 kyr the North Atlantic warming and Northern Hemisphere deglaciation. A more important result of our study is that there have been several periods during the last glacial of large isotopic anomalies in the surface waters of the Antarctic polar front, particularly between 35 and 17 kyr BP. These anomalies are due to a large input of melt water during periods when the periphery of the Antarctic ice sheet was rapidly eroded. We propose that these events occurred during maximum extension of the marine-based ice sheets, and were due to the rapid erosion of the periphery of the Antarctic ice sheet by ice streams, flowing towards the ice shelves. When most of the ice available for feeding the coastal ice streams had been eroded, the ice shelves were destroyed, and iceberg calving decreased dramatically, until a new equilibrium ice-sheet profile was attained.

This work was supported by Programme National d'Etude de la Dynamique du Climat, Mission Scientifique des Terres Australes et Antarctiques Françaises, Centre National de la Recherche Scientifique, Commissariat à l'Energie Atomique, Université de Bordeaux I and Institut Géologique du Bassin d'Aquitaine. Sediment samples came from the Centre des Faibles Radioactivités and the Lamont-Doherty Geological Observatory core collections (the latter is supported by NSF). We acknowledge support from J. Antignac, H. Leclaire and B. Le Coat for the isotopic analyses, L. Burckle, J. Jouzel, C. Lorius, J. Moyes and J. L. Turon for useful discussions, M. Fontugne as chief scientist of the cruise APSARA 2, and captains, officers, and crews of M.S. *Marion Dufresne* for their work during the different cruises in the Southern Ocean. Drs S. Calvert, J. Hays and N. Shackleton reviewed and improved the manuscript. C. F. R. contribution 782.

Received 1 April; accepted 19 June 1986.

1. CLIMAP Science 191, 1131–1137 (1976).
2. Knox, F. & McElroy, M. E. *J. Geophys. Res.* 89, 4629–4637 (1984).

3. Hollin, J. T. *J. Glaciol.* **4**, 173-175 (1962).
4. Imbrie, J. & Kipp, N. J. in *The Late Cenozoic Ages* (ed. Turekian, K. K.) 71-79 (Yale University Press, New Haven, 1971).
5. Pichon, J. J., Labeyrie, M. & Labeyrie, L. D. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* (submitted).
6. Emiliani, C. *J. Geol.* **74**, 109-124 (1966).
7. Gordon, A. I. in *Antarctic Oceanology 1* (*Antarctic Research Series Vol. 15*) (ed. Reid, J. L.) 169-203 (American Geophysical Union, Washington, 1971).
8. Hays, J. D., Lozano, J. A., Shackleton, N. & Irving, G. *Mem. Geol. Soc. Am.* **145**, 337-372 (1976).
9. Cooke, D. W. & Hays, J. D. *Proc. 3rd Symp. Antarct. Geol. Geophys.* (ed. Craddock, C.) 1017-1025 (University of Wisconsin Press, Madison, 1982).
10. Duplessy, J. C., Moyes, J. & Pujol, C. *Nature* **286**, 279-282 (1980).
11. Labeyrie, L. D. & Duplessy, J. C. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **50**, 217-240 (1985).
12. Caulet, J. P. *Bull. Soc. Géol. Fr.* **24**, 555-562 (1982).
13. Imbrie, J. *et al.* in *Milankovitch and Climate* (eds Berger, A. L., Imbrie, J., Hays, J., Kukla, G. & Saltzman, B.) 269-306 (Reidel, New York, 1984).
14. Duplessy, J. C., Labeyrie, L. D., & Shackleton, N. J. *Eos* **66**, 292 (1985).
15. Labeyrie, L. D., Juillet, A. & Duplessy, J. C. in *Proc. 7th Diatom Symp. 1982* (ed. Cramer, J.) 477-491 (Gantner, Vaduz, 1984).
16. Duplessy, J. C. *et al.* *Nature* **320**, 350-352 (1986).
17. Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* **194**, 1121-1132 (1976).
18. Kellogg, T. B., Duplessy, J. C. & Shackleton, N. J. *Boreas* **7**, 61-71 (1978).
19. Reynolds, L. A. & Thunell, R. C. *Micropaleontology* **32**, 1-18 (1986).
20. MacIsaac, J. J., Dugdale, R. C., Barber, R. T., Blasco, D. & Packard, T. T. *Deep Sea Res.* **32**, 503-529 (1985).
21. Shackleton, N. J. in *Variations du Climat au Cours du Pleistocene*, 203-209 (Centre National de la Recherche Scientifique, Gif, 1974).
22. Labeyrie, L. D. *et al.* *Geol. Soc. Am. Abstr. Progr.* **17**, 637 (1985).
23. Craig, H. & Gordon, L. in *Stable Isotopes in Oceanographic Studies and Paleotemperatures*, 9-131 (Consiglio Nazionale delle Ricerche, Pisa, 1965).
24. Duplessy, J. C. *C. R. hebdomadaire Séances Acad. Sci., Paris* **271**, 1075-1078 (1970).
25. Weiss, R. F., Ostlund, H. G. & Craig, H. *Deep Sea Res.* **26A**, 1093-1120 (1979).
26. Cooke, D. W. *thesis*, Columbia Univ. (1978).
27. Smith, D. G., Ledbetter, M. T. & Cieselski, P. F. *Mar. Geol.* **53**, 291-312 (1983).
28. Keany, J., Ledbetter, M., Watkins, N. & Huang, T. C. *Geol. Soc. Am. Bull.* **87**, 873-882 (1976).
29. Morley, J. J. & Hays, J. *Earth planet. Sci. Lett.* **66**, 63-72 (1983).
30. Burckle, L. H. *Mar. Micropaleontol.* **9**, 241-261 (1984).
31. Jacobs, S. S., Gordon, A. L. & Amos, A. F. *Nature* **277**, 469-471 (1979).
32. Oriheim, O. *U.S. Dep. Energy Rep. No. DOE/EV/602351*, 210-215 (1984).
33. Lorius, C., Merlivat, L., Jouzel, J. & Pourchet, M. *Nature* **280**, 644-648 (1979).
34. Lorius, C. *et al.* *Nature* **316**, 591-596 (1985).
35. Yiou, F., Raisbeck, G. M., Boursles, D., Lorius, C. & Barkov, N. I. *Nature* **316**, 616-617 (1985).
36. Raynaud, D. & Whillans, I. M. *Ann. Glaciol.* **3**, 269-273 (1982).
37. Denton, G. H. & Hughes, T. J. *Quat. Res.* **20**, 125-144 (1983).
38. Hughes, T. J. *et al.* in *The Last Great Ice Sheets* (eds Denton, G. H. & Hughes, T. J.) 263-317 (Wiley, New York, 1981).
39. Sugden, D. E. & Clapperton, C. M. *Quat. Res.* **7**, 268-282 (1977).
40. Burkle, L. H. & Cooke, D. W. *Micropaleontology* **29**, 6-10 (1983).
41. Neshiba, S. *Nature* **267**, 507-508 (1977).
42. Weertman, J. *Nature* **260**, 284-286 (1976).
43. Thomas, R. H. *Nature* **259**, 180-183 (1976).
44. Chappell, J. *Search* **14**, 99-101 (1983).
45. Wefer, G. *et al.* *Nature* **299**, 145-147 (1982).
46. Dietrich, G. *Atlas of the Hydrography of the Northern North Atlantic Ocean* (Conseil International pour l'Exploration de la Mer, Charlottenlund Slot, 1969).
47. Wyrki, K. *Oceanographic Atlas of the International Indian Ocean Expedition* (National Science Foundation, Washington, 1971).
48. Duplessy, J. C. *et al.* *Quat. Res.* **21**, 225-243 (1984).
49. Tchernia, P. & Jeannin, P. F. *Quelques Aspects de la Circulation Océanique Antarctique Révélés par l'Observation de la Dérive d'Icebergs* (Centre National d'Etudes Spatiales, Toulouse, 1983).

Generation of single-stranded T-DNA molecules during the initial stages of T-DNA transfer from *Agrobacterium tumefaciens* to plant cells

Scott E. Stachel*, Benedikt Timmerman & Patricia Zambryski*

Laboratorium voor Genetica, Rijksuniversiteit Gent, B-9000 Gent, Belgium

Activation of the T-DNA transfer process of Agrobacterium by the plant signal molecule acetosyringone generates a single-stranded, unipolar, linear T-DNA molecule (T-strand)—a potential conjugative intermediate in the transfer of the T-DNA to plant cells. Acetosyringone induction also leads to S₁ nuclease-sensitive sites at the Ti plasmid T-DNA borders, and other molecular changes associated with the Ti plasmid T-DNA sequences, which may correspond to specific steps of T-strand synthesis.

DURING the genetic transformation of plant cells by the soil pathogen *Agrobacterium tumefaciens* (reviewed in ref. 1), a specific segment of DNA, the T-DNA, is recognized in and mobilized from the large (>200 kilobase pairs; kbp) Ti plasmid of the bacterium, transferred across the cell walls of the bacterium and plant cell, and integrated as an unaltered fragment into the plant nuclear genome. Analysis of the transfer process has focused on what defines the T-DNA; the genetic requirements for transfer other than the T-DNA; and the mechanism of transfer.

In the Ti plasmid the T-DNA is bounded by essentially identical 25-base pair (bp) direct repeats²⁻⁵. These sequences define the T-DNA, for any DNA, and only DNA, located between T-DNA borders is efficiently transferred and integrated⁶⁻⁹. The T-DNA transfer process is directed by the products of the Ti plasmid virulence (*vir*) and chromosomal virulence (*chv*) loci (reviewed in refs 10, 11). Whereas *chv*

expression is constitutive, *vir* expression is tightly regulated^{12,13}, and its activation initiates the transfer process. This activation is mediated by specific phenolic compounds present in the exudates of wounded and actively metabolizing plant cells. One such compound is acetosyringone (AS; 4-acetyl-2,6-dimethoxyphenol)¹⁴.

Genetic analyses of the 25-bp sequences have indicated that they are polar in function^{15,16}. While deletion of the left border repeat has no significant effect on pathogenicity¹⁷, deletion of the right repeat totally abolishes it^{15,16,18}. Furthermore, when the orientation of the right border is reversed with regard to its natural orientation on the Ti plasmid, the efficient transfer and/or integration of the T-DNA sequences is greatly attenuated. These results indicate that T-DNA transfer may occur in a rightward to leftward fashion, determined by the orientation of the 25-bp border repeats, and suggest that transfer might be via a conjugative mechanism¹⁵.

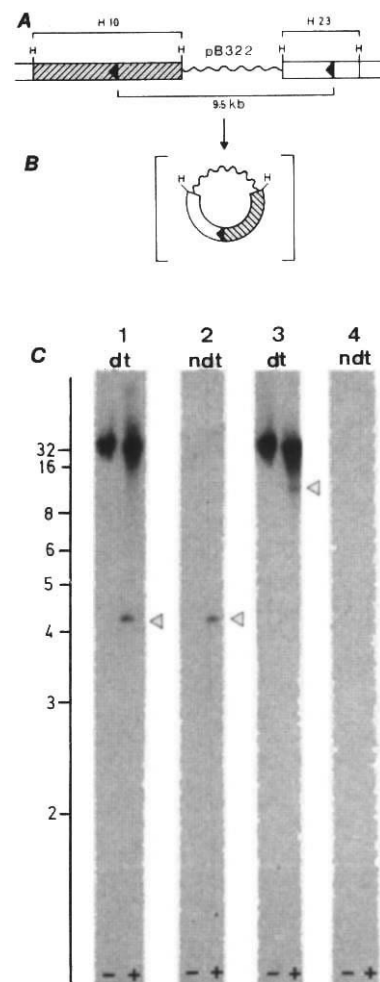
Here we use hybridization techniques to directly identify and characterize novel structures associated with the T-DNA and its border sequences in *Agrobacterium* following induction of *vir* gene expression with AS. A variation of the Southern blotting transfer procedure is used to distinguish between single-stranded (ss) and double-stranded (ds) DNA molecules present in total

* Present address: Institute of Cancer Research and Howard Hughes Medical Institute, College of Physicians and Surgeons, Columbia University, New York 10032, USA (S.E.S.); Division of Molecular Plant Biology, Hilgard Hall, University of California, Berkeley, California 94720, USA (P.Z.).

Fig. 1 Hybridization analysis of total DNA prepared from AS-induced *Agrobacterium*. **A**, T-DNA region of the pGV3850 Ti plasmid⁶. The interior portion of the T-DNA of pTiC58 has been replaced with pBR322 (wavy line). *Hind*III fragments 10 (hatched) and 23 (white) from pTiC58 carrying the right and left 25-bp T-DNA border sequences respectively (black arrows); H, *Hind*III restriction site. **B**, Structure of the 9.5-kbp ds T-DNA circle molecule isolated after transformation of *E. coli* with undigested total DNA from AS-induced pGV3850 *Agrobacterium*. **C**, Southern blot hybridization analyses of *Agrobacterium* DNA. Total DNA prepared from uninduced (–) and AS-induced (+) 3850 *agrobacteria* was analysed by hybridization against nick-translated ³²P-labelled pBR322 probe. Lane 1, denatured transfer (dt) of untreated DNA; lane 2, non-denatured transfer (ndt) of untreated DNA; lane 3, denatured transfer of total DNA digested with S₁ nuclease; lane 4, non-denatured transfer of S₁-treated DNA. Scaling on left is in kbp. Arrows indicate novel signals observed in the AS-induced lanes. Note that all of these novel signals are also observed, albeit at a lower level, with DNA prepared from *Agrobacterium* co-cultivated with *Nicotiana tabacum* protoplast cells (data not shown).

Methods. An overnight culture of 3850 *Agrobacterium* grown in YEB liquid medium was resuspended in M5SP medium¹⁴ at 0.05 A₆₀₀ units ml^{–1} and grown at 28 °C with high aeration. After 5 h of preincubation growth, acetosyringone (Janssen) was added to half of the M5SP culture at 100 µM, a concentration that is non-limiting for induction of the Ti-plasmid *vir* genes¹⁴. The uninduced and AS-induced cultures were grown for another 12–18 h and a parallel culture of the *virB::lac* strain A348(pSM30) (refs 10, 14) was used to monitor *vir* induction; under these conditions the cultures undergo two to three doublings, and >100-fold increases in β-galactosidase activity are observed in A348(pSM30). The bacteria were collected by centrifugation and total DNA prepared as described previously³³. Briefly, the pellet from 5 ml of cells is lysed in 200 µl TE (50 mM Tris, 20 mM Na₂-EDTA, pH 8.0), 100 µl 5% sodium sarkosyl and 100 µl pronase (10 mg ml^{–1}) for 45 min at 37 °C. The lysate is then vortexed for 15 s (light-shear), extracted twice with phenol and twice with chloroform, and the DNA recovered by EtOH precipitation. *E. coli* was transformed with uninduced and AS-induced DNA as described previously¹⁹, and the AS-induced preparation was determined to give transformants carrying the ds T-DNA circle molecule diagrammed in **B**. Aliquots (1 µg) of total *Agrobacterium* DNA (either untreated or S₁-digested) were then electrophoresed in 0.9% TBE agarose gels containing 0.5 µg ml^{–1} EtBr, transferred to nitrocellulose in 10× SSC, and analysed by Southern blot filter hybridization. Two different transfer conditions have been used, denatured and non-denatured. For denatured transfer, the agarose gel is soaked in denaturing solution for 60 min followed by neutralizing solution for 60 min before capillary blotting. For non-denatured transfer, the gel is soaked in H₂O for 10 min, then 10× SSC for 10 min, before blotting. For the S₁ nuclease digestions, 1 µg total DNA in 200 µl S₁ digestion buffer was incubated with 50 U S₁ nuclease (Boehringer Mannheim) for 30 min at 20 °C. The reaction is terminated by adding 20 µl 10× S₁ termination buffer, followed by phenol extraction and EtOH precipitation. Unless otherwise specified, all buffers and conditions used here and in

Figs 2–5 are according to Maniatis *et al.*³⁴.



DNA prepared from these cells. Evidence is presented for (1) free ss T-DNA molecules (T-strands) whose polarity corresponds to that of the T-DNA borders; (2) ss endonuclease-sensitive structures associated with the Ti plasmid T-DNA borders (S₁ border sites); and (3) molecular alterations associated with the internal sequences of the Ti plasmid T-region (T-region structures). We discuss the potential role of each of these AS-induced T-DNA homologous molecules in the transfer of the T-DNA to the plant cell.

Free T-DNAs in AS-induced cells

For these studies we used *Agrobacterium* carrying Ti plasmid pGV3850 (ref. 6). pGV3850 has been derived from the nopaline C58 Ti plasmid, and the structure of its minimal T-DNA region is shown in Fig. 1A. Strain 3850 has been used previously to isolate and identify a 9.5 kbp ds T-DNA circle molecule (Fig. 1B) following transformation of *Escherichia coli* with total DNA prepared from plant¹⁹ or AS-induced *Agrobacterium*¹⁴; this T-DNA circle has been proposed as a candidate for the T-DNA molecule that is transferred to the plant cell. Our initial experiments aimed to identify the presence of this molecule directly in *Agrobacterium*.

Total DNA was prepared from AS-induced cells and found to produce ds T-DNA circles in *E. coli*; this AS-induced DNA, along with total DNA prepared from uninduced cells, was fractionated by agarose gel electrophoresis and transferred to a nitrocellulose filter following gel denaturation (normal Southern transfer procedure, see below). Two T-DNA homologous signals are observed in Fig. 1C, lane 1. Since the DNA is undigested, the upper signal represents Ti-linked T-DNA sequences, while

the lower signal represents free T-DNA sequences. This novel signal is specific to the AS-induced DNA, and thus corresponds to a Ti-independent T-DNA molecule whose synthesis is the result of the AS-induced activation of the pGV3850 *vir* loci.

The free T-DNA signal does not migrate as a 9.5-kbp ds DNA molecule (supercoiled, relaxed-circular or linear), but instead migrates at a size corresponding to a ds linear molecule of 4.4 kbp. Thus, by hybridization we find no evidence for ds T-DNA circles in AS-induced bacteria. This result is not totally unexpected because the frequency of recovery of these molecules in *E. coli* is low. We obtain on average 50 ds circle transformants per µg AS-induced DNA (at a transformation efficiency of 10⁷ transformants per µg supercoiled pBR322 plasmid DNA). Since 1 µg of total *Agrobacterium* DNA contains ~1.7 ng of pGV3850 T-DNA¹⁹, at most only one in every 100 AS-induced cells harbours a ds T-DNA circle. That these molecules are not detected by hybridization indicates that the transformation results are representative of their actual concentration in the AS-induced DNA. We discuss below a model for how ds circle molecules might be generated at a low frequency as a result of AS induction (Fig. 6).

Free T-DNAs are single stranded

Since the size of the AS-induced free T-DNA molecule is about half that of the T-region of pGV3850 (Fig. 1A), it may be a single-strand copy of the pGV3850 T-DNA. The following experiments demonstrate that the free T-DNA molecule has properties characteristic of ss DNA. Evidence is also given for other novel T-DNA-homologous molecules present in the AS-induced bacteria; in contrast to the free T-DNA, these structures

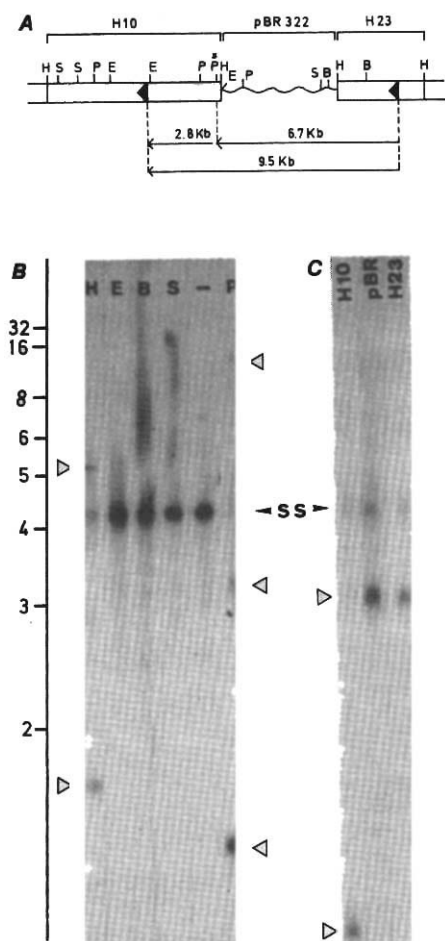


Fig. 2 Restriction endonuclease treatment of AS-induced DNA. **A**, Restriction map of the T-DNA region of pGV3850. Black arrows, T-DNA border sequences; the 9.5-kb arrow designates the ss T-DNA molecule, and the 2.8-kb and 6.7-kb arrows correspond to the ss fragments that would be produced by *PstI* cleavage of the ss T-DNA molecule at the *PstI** site. **B**, Aliquots (1 µg) of AS-induced DNA were digested to completion with five restriction endonucleases, fractionated on 0.9% agarose, transferred under non-denaturing conditions to nitrocellulose, and analysed by hybridization against *HindIII* fragment 10/pBR322 probe. H, *HindIII*; E, *EcoRI*; B, *BamHI*; S, *Sall*; P, *PstI*; -, undigested; ss, free T-DNA molecule; open arrows, unexpected fragments produced by *HindIII* or *PstI* digestion of the AS-induced DNA. **C**, Three equivalent non-denatured transfers of *PstI*-digested AS-induced DNA were hybridized against *HindIII* fragment 10 probe (H10); pBR322 probe (pBR); and *HindIII* fragment 23 probe (H23). Scaling is different from that in **B**, such that the signals marked with triangles correspond to ss DNA molecules of ~6.6 and 2.8 kb, respectively. Note that different preparations of AS-induced DNA were used in the experiments of **B** and **C**. The ~9.2-kbp signal observed in the *PstI* digest in **B** is not observed in **C**; this difference may reflect the detection of different levels of AS-induced events in the two preparations.

are still linked to the Ti plasmid, as they are only detected after enzymatic digestion.

Transfer assay for ss T-DNAs. DNA binds to nitrocellulose only if it is single-stranded²⁰ or associated with protein²¹. In the Southern blotting technique, ds DNA fragments fractionated in an agarose gel must be denatured before being transferred to nitrocellulose²⁰. Thus, ss (and protein-associated) DNA molecules can be easily distinguished from ds molecules by comparing their transfer to nitrocellulose from agarose gels treated with or without NaOH before transfer. Figure 1C, lanes 1 and 2, show the denatured and non-denatured transfers of

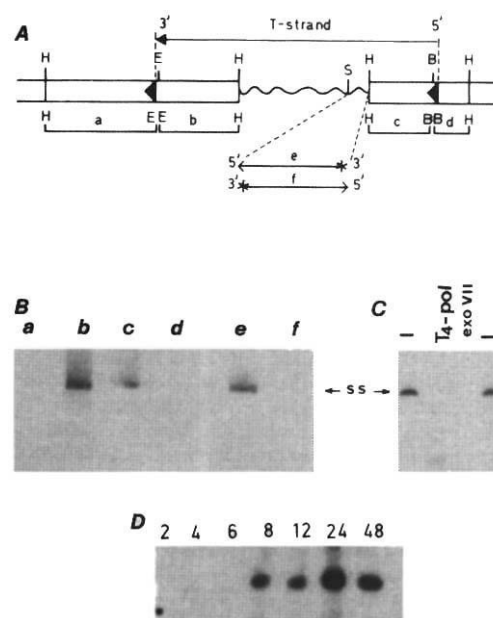


Fig. 3 Molecular characterization of ss T-DNA molecules. **A**, Six probes used to analyse the sequence content and strandedness of the ss T-DNA molecule. **B**, *BclI*; E, *EcoRI*; H, *HindIII*. The *EcoRI* site of probes *a* and *b* falls 55 bases inside the left T-DNA border, and the *BclI* site of probes *c* and *d* falls 42 bases inside the right border.

Methods. Fragments *a*, *b*, *c* and *d* were gel-purified twice and labelled by nick translation. Probes *e* and *f* are 3'-end-labelled probes where the position of the label is marked by an asterisk, and correspond to the upper and lower strands of the pGV3850 T-DNA, respectively. For probe *e*, pBR322 was digested with *HindIII*, 3'-end-labelled at the *HindIII* site by Klenow fill-in, recut with *Sall*, and the resultant 622-bp *HindIII/Sall* fragment was gel-purified. For probe *f*, the order of the restriction digests was reversed to give a 622-bp *Sall/HindIII* fragment, 3'-end-labelled at the *Sall* site. The polarity of the ss T-DNA molecule, or T-strand, is shown at the top of the figure. **B**, Sequence content and strandedness of the ss T-DNA. Six equivalent strips of a non-denatured transfer of undigested AS-induced DNA hybridized against the six probes shown in **A**. ss, Single-stranded T-DNA signal. **C**, Exonuclease sensitivity of the ss T-DNA. The non-denatured transfer of untreated (-), T4 polymerase-digested (T4-pol), and exonuclease VII-digested (ExoVII) AS-induced DNA was hybridized against nick-translated pBR322 probe. For exonuclease digestion, 1-µg aliquots of AS-induced DNA were incubated with 2.5 U T4 polymerase (Anglian Biotechnology Ltd) or 1.1 U exonuclease VII (Gibco-BRL), respectively, for 30 min at 37 °C. Each reaction was carried out in 15 µl in the enzyme buffer recommended by the manufacturer. Note that under identical conditions these treatments have no effect on M13 ss circular phage DNA; also, the ss T-DNA is fully degraded after 90 min digestion with T4 polymerase (data not shown). **D**, Kinetics of synthesis of the ss T-DNA molecule. A 50-ml culture of pGV3850 *Agrobacterium* pregrown for 5 h was then induced with AS at 100 µM. At 2, 4, 6, 8, 12, 24 and 48 h after the start of induction, 5-ml aliquots were removed, frozen at -70 °C, and total DNA was then prepared. These DNA samples were fractionated on agarose and transferred to nitrocellulose without gel denaturation, and the amount of ss T-DNA molecules present in the AS-induced cells for each time point was assessed by hybridization against nick-translated pBR322 probe. Note that by EtBR staining the amount of DNA in the 24-h lane was about double that for the other lanes.

uninduced and AS-induced total DNA. Of the two T-DNA-homologous signals present in the AS-induced DNA sample, only the lower signal, which represents the free T-DNA molecule, transfers without gel denaturation (Fig. 1C, lane 2). This free T-DNA must be ss DNA because the AS-induced total DNA has been fully deproteinized during its preparation (Fig. 1

legend) and RNase digestion of the DNA before hybridization analysis does not affect the ss signal (data not shown).

Nuclease sensitivity of ss T-DNAs. S_1 nuclease is a ss-specific endonuclease²². In Fig. 1C, the denatured (lane 3) and non-denatured (lane 4) transfers of S_1 nuclease-digested samples of uninduced and AS-induced total DNA demonstrate that the free T-DNA molecule is fully degraded by S_1 nuclease. Interestingly, S_1 treatment also appears to release the T-region from the AS-induced Ti plasmid. A novel T-DNA homologous signal is observed in the S_1 -treated AS-induced sample in the denatured transfer (Fig. 1C, lane 3), and this signal corresponds to a ds molecule, which migrates as a linear fragment of 9.5 kbp, the precise size of the pGV3850 T-DNA from its left to right T-DNA borders (Fig. 1A). Thus, AS induction also results in a Ti plasmid whose T-DNA borders are sensitive to cleavage by S_1 nuclease (see Figs 4 and 5).

The activity of type II restriction endonucleases is limited primarily to duplex DNA. Aliquots of AS-induced DNA were digested to completion with *Bam*HI, *Eco*RI, *Hind*III, *Pst*I and *Sal*I (each of these enzymes cleaves within the T-DNA region of pGV3850; Fig. 2A), and hybridized to the *Hind*III fragment 10/pBR probe following non-denatured transfer (Fig. 2B). Because the ss signal is present in all the digests (except for *Pst*I), the free T-DNA is generally resistant to restriction digestion.

The *Hind*III and *Pst*I results are more complex. The ss signal is weak in the *Hind*III lane and absent in the *Pst*I lane; also, several novel signals are observed in these lanes. Several restriction enzymes have been shown to restrict ss DNA²³, and the sensitivity of the free T-DNA to *Hind*III and *Pst*I may be an example of this effect. Interestingly, fragments both smaller and larger than the ss T-DNA molecule are produced by *Hind*III and *Pst*I digestion (Fig. 2B). While the smaller fragments can be cleavage products of the ss T-DNA, the larger fragments must be derived from a large molecule, presumably the AS-induced Ti plasmid. Thus, AS induction also results in a Ti plasmid molecule whose T-region carries ss sequences (see Fig. 5).

Characterization of ss T-DNAs

ss T-DNAs are full-length and unipolar. Six identical strips of a non-denatured transfer of undigested AS-induced DNA were hybridized separately with the six probes labelled *a* to *f* in Fig. 3A. Probes *a* and *d* correspond to pGV3850 sequences just outside the left and right T-DNA border repeats; probes *b* and *c* correspond to T-region sequences just within these repeats; and probes *e* and *f* correspond to the upper and lower strands of the pGV3850 T-DNA region, respectively. Figure 3b shows that the ss T-DNA molecule hybridizes only to probes *b*, *c* and *e*. Thus, it is composed only of sequences located internal to the T-DNA borders, and which correspond to the lower strand of the pGV3850 T-region. We designate this unipolar ss T-DNA molecule as the T-strand.

T-strands are linear. The T-strand is sensitive to both exonuclease VII (3'→5' and 5'→3' activities specific for ss DNA²⁴) and T4 polymerase (3'→5' exonuclease that accepts both ss and ds DNA as substrate²⁵) (Fig. 3C). Thus, the T-strand, as isolated, must be a linear ss DNA molecule whose 3' terminus (and perhaps also 5' terminus) is available to exonuclease digestion.

The *Pst*I cleavage products of the T-strand indicate that its 5' and 3' ends map to the right and left T-DNA borders, respectively. Two fragments of ~3.3 and 1.4 kbp (~ss lengths 6.7 and 2.8 kbp) are detected. The 6.6-kbp fragment is homologous to pBR322 and *Hind*III fragment 23, while the 2.8-kbp fragment is homologous only to *Hind*III fragment 10 (Fig. 2C). Cleavage of a circular T-strand molecule at any (or all) of its three *Pst*I sites would yield fragments of different sizes and sequence content from those observed. The simplest explanation for the results of Fig. 2C is that the T-strand is a linear ss molecule, cleaved by *Pst*I principally at its middle *Pst*I site (just within

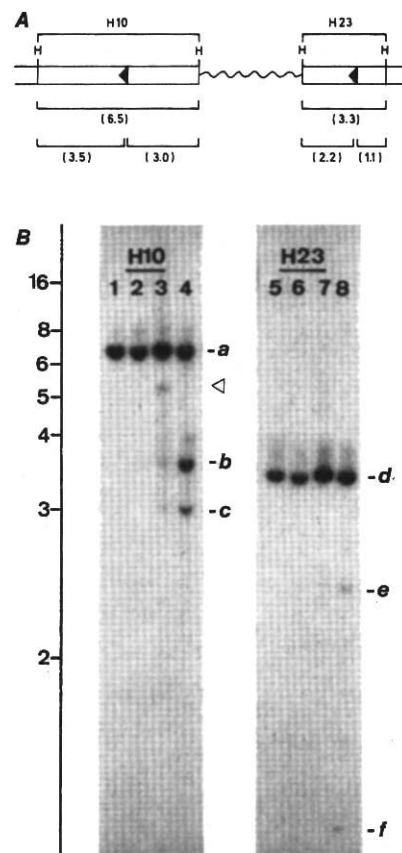


Fig. 4 Identification of AS-induced S_1 -sensitive T-DNA border structures. **A**, *Hind*III fragments produced by ds cleavages of the pGV3850 T-DNA borders. Fragments *b* and *c* result from cleavage of the left border repeat carried by *Hind*III fragment 10 (*a*). Fragments *e* and *f* result from cleavage of the right T-DNA border carried by *Hind*III fragment 23 (*d*). Fragment sizes are in kbp. Black arrows, T-DNA border sequences; H, *Hind*III. **B**, Uninduced and AS-induced DNA was digested with *Hind*III, and half of each sample was treated further with S_1 nuclease. The four samples were analysed on two denatured transfers hybridized against nick-translated *Hind*III fragment 10 (lanes 1-4) or *Hind*III fragment 23 probe (lanes 5-8). Lanes 1, 5, uninduced DNA; lanes 2, 6, uninduced DNA plus S_1 treatment; lanes 3, 7, AS-induced DNA; lanes 4, 8, AS-induced DNA plus S_1 treatment. *a-f* refer to the fragments shown in **A**. The open arrow indicates a novel S_1 -sensitive fragment of ~5.2 kbp. Scaling is in kbp. Note that the low level of the border cleavage fragments (*b*, *c*, *e*, *f*) seen in the absence of S_1 treatment (lane 3, and lane 7 after long exposure) may reflect mechanical breakage of the AS-induced S_1 -sensitive border structures during sample preparation.

the right end of its *Hind*III fragment 10 sequences), to produce a 6.7-kbp 5' and a 2.8-kbp 3' fragment.

T-strand synthesis is limited. Using a probe homologous to both strands of the T-DNA, the intensity of the T-strand signal is ~5-10-fold less than the Ti-plasmid T-region signal (Fig. 1C, lane 1; also Fig. 5). As the T-strand corresponds to one strand of the T-region, and the copy number of the Ti plasmid is about two (and assuming no specific loss of T-strands during DNA preparation), we estimate that on average each AS-induced cell carries 0.4-0.8 T-strand molecules. As T-DNA homologous molecules are not found in the culture medium (data not shown), this low copy number is probably not due to export of the T-strand. Figure 3D demonstrates that the relative amount of T-strands begins to plateau within 8 h after the start of AS induction, and shows only a doubling during an additional 40 hours of induction. Thus, T-strand synthesis is a limited process.

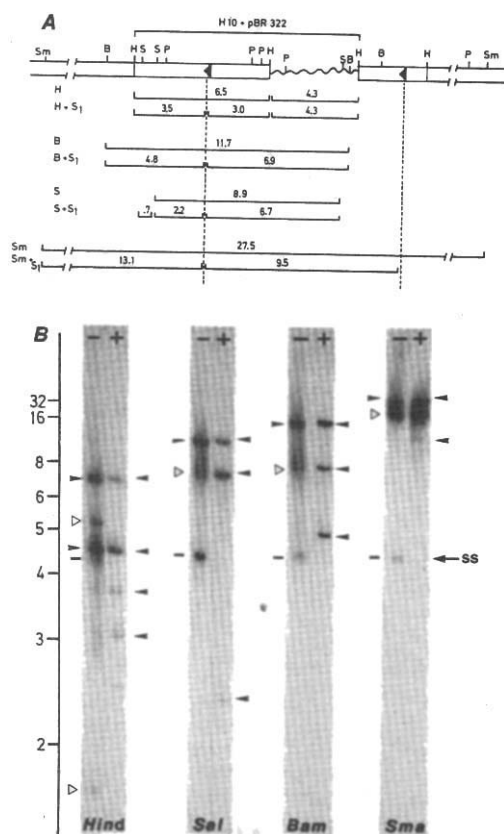


Fig. 5 Hybridization analysis of T-region intermediate structures. **A**, Restriction map of the left portion of the pGV3850 T-DNA region and border cleavage fragments predicted to arise after *S*₁ treatment of AS-induced DNA. Sizes are given in kbp. **B**, *Bam*HI; *H*, *Hind*III; *P*, *Pst*I; *S*, *Sal*I; *Sm*, *Sma*I; *S*₁, *S*₁ nuclease. The black arrows and dotted lines indicate the positions of the T-DNA borders. **B**, Denatured transfers of *Hind*III, *Bam*HI, *Sal*I and *Sma*I digests of AS-induced DNA treated (+) and untreated (-) with *S*₁ nuclease were hybridized against nick-translated *Hind*III fragment 10/pBR322 probe (**A**). Scaling on left is in kbp. The single-stranded T-strand is indicated by ss and the bar adjacent to the digests. Solid arrows indicate the expected fragments before and after *S*₁ treatment and correspond in size to the fragments shown in **A**. The open arrows indicate unexpected AS-induced fragments which correspond to novel T-DNA homologous structures linked to pGV3850. Note that other unexpected signals are observed when the filter is reprobed with *Hind*III fragment 23 sequences, and that all of the unexpected fragments are specific to AS-induced DNA and have been observed with DNA prepared from at least two independent AS-induced bacterial cultures (data not shown).

AS-induced alterations of the Ti plasmid

Since the T-strand must be produced from the Ti-plasmid T-region, several molecular reactions involving these sequences should occur in bacterial cells that are actively carrying out different steps of T-strand synthesis, or other early steps of T-DNA transfer. To gain insight into these reactions, the novel T-DNA homologous hybridization signals that result from fragmentation of the AS-induced Ti plasmid with restriction enzymes and *S*₁ nuclease are characterized below. These data provide evidence for specific reactions associated with the Ti-plasmid T-DNA border repeats and with the internal sequences of the T-region.

***S*₁-sensitive T-DNA border sites.** Figure 4 demonstrates that AS induction leads to the generation of Ti-plasmid T-DNA border structures which are sensitive to cleavage by *S*₁ nuclease. Uninduced and AS-induced DNA was restricted with *Hind*III, half of each sample was treated further with *S*₁ nuclease, and the

fragments homologous to the left and right border regions of pGV3850 were detected by hybridization. The AS-induced DNA lanes contain novel signals whose sizes correspond exactly to the fragments that would be produced by ds cleavages at the left and right T-DNA border sequences of pGV3850 (Fig. 4A), and the intensities of these signals are greatly increased by *S*₁ digestion of the DNA. We designate these *S*₁-sensitive structures *S*₁ border sites.

The relative intensities of the signals which correspond to the *S*₁ border cleavage fragments show that ~30% of the left T-DNA borders and ~10% of the right T-DNA borders of the AS-induced Ti plasmid population are *S*₁-sensitive. Thus, the generation of *S*₁ border sites is a frequent event. Furthermore, since a 9.5-kbp ds T-DNA fragment is produced by *S*₁ digestion of AS-induced DNA that is unrestricted (Fig. 1C, lane 3), or restricted with *Sma*I (which does not cleave within the pGV3850 T-region (Fig. 5)), *S*₁ border sites can simultaneously occur at both the right and left border repeats on a single Ti plasmid. We note that incubation of the uncut AS-induced DNA at 65 °C does not result in the release of the 9.5-kbp T-DNA fragment from the AS-induced Ti plasmid (data not shown); that is, the *S*₁ border site does not correspond to a ds staggered cleavage of the T-DNA border sequence. Other experiments have shown that the *S*₁-sensitive structure corresponds to an AS-induced ss endonucleolytic cleavage in the bottom strand of the 25-bp border sequence (K. Wang and M. Van Montagu, in preparation); *S*₁ is known to cleave opposite such structures^{26,27}.

Other AS-induced T-region structures. Figure 5 demonstrates that other novel Ti-plasmid-linked T-region structures are also present in the AS-induced cell population. Three different classes of T-DNA-homologous signals are observed when AS-induced DNA is digested with various restriction enzymes and *S*₁ nuclease (Fig. 5B): those that correspond to the ss T-DNA molecule (black bars, -*S*₁ lanes); those that correspond to the predicted border cleavage fragments (Fig. 5A) produced by *S*₁ digestion of the AS-induced *S*₁ border sites (solid arrows, +*S*₁ lanes); and those that correspond to novel unexpected fragments (open triangles) derived from pGV3850 Ti plasmids whose internal T-region sequences have been altered as a result of AS induction.

For example, the 5.2-kbp fragment present in the -*S*₁ *Hind*III sample (Fig. 5B; also Figs 2B and 4B) is *S*₁-sensitive, binds to nitrocellulose, specifically hybridizes to *Hind*III fragment 10 (and not to pBR322 or *Hind*III fragment 23; data not shown), and is ~1.3 kbp smaller than *Hind*III fragment 10 (Fig. 5A). Thus, this novel *Hind*III fragment must be partially single-stranded and derived from an AS-induced Ti plasmid whose T-region is partially single-stranded or associated with ss T-DNA sequences, as in a D-loop structure. More perplexing fragments are also detected: the ~15-kbp *S*₁-insensitive *Sma*I signal corresponds to a ds DNA fragment which is ~12 kbp smaller than the *Sma*I fragment that covers the T-region of the uninduced Ti plasmid; also, in the -*S*₁ *Bam*HI and *Sal*I lanes, signals are observed whose sizes correspond to the *S*₁ cleavage fragments of these digests (Fig. 5A) which are internal, but not external, to the left T-DNA border. While the present data do not allow the precise identification of the T-region structure(s) to which these unexpected fragments correspond, they illustrate the complexity of the molecular reactions associated with the Ti-plasmid T-DNA sequences which occur in AS-induced *Agrobacterium*.

Discussion

Agrobacterium tumefaciens transfers its Ti-plasmid T-DNA to plant cells, and this process is activated by the induction of the Ti-plasmid virulence genes with the plant phenolic compound AS. We show that AS induction results in the generation of several novel T-DNA homologous molecules in *Agrobacterium*: a linear ss molecule, the T-strand, which corresponds to the lower strand of the Ti-plasmid T-region; and Ti-plasmid molecules whose T-DNA border repeats are sensitive to cleavage

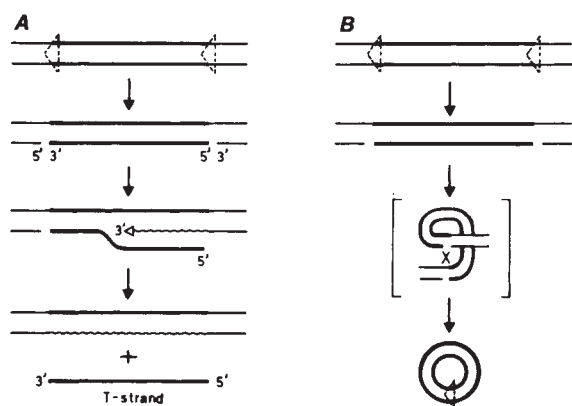


Fig. 6 Proposed reactions associated with the Ti-plasmid T-region in AS-induced cells. **A**, Model for the generation of the T-strand; **B**, model for the generation of a ds T-DNA circle molecule. Thick lines represent T-region sequences, thin lines represent adjoining Ti-plasmid sequences external to the 25-bp T-DNA border repeats, indicated by the dashed triangles. ss breaks in the bottom strand of the border repeats correspond to S_1 border sites, and the wavy line corresponds to a newly synthesized bottom strand of the T-region. This model implies that ss molecules are not generated leftward of the left T-DNA border; while we have not observed such molecules, they might not have been detected if they are heterogeneous in length.

by S_1 nuclease (S_1 border sites), and whose internal sequences have been altered (T-region structures). The specific properties of these AS-induced molecules allows the formulation of mechanistic models for the generation of the T-strand, and for its potential transfer to the plant cell.

Figure 6A presents a model for the generation of the T-strand molecule. First, ss endonucleolytic cleavages occur within the left and right T-DNA border sequences on the Ti plasmid. These cleavages correspond to S_1 border sites and provide free 3' OH groups from which DNA synthesis can be primed. Second, using the top strand of the T-region as template, DNA synthesis initiates at the right border cleavage site, and proceeds unidirectionally across the T-region. This synthesis displaces the bottom strand of the T-region and produces a transitory triple-stranded structure which may correspond to one of the AS-induced T-region structures that we have detected. Third, DNA synthesis terminates when it encounters the left border cleavage site, and the displaced bottom strand is released from the Ti plasmid, as the free T-strand ss molecule. Alternatively, a 5'→3' helicase activity unwinds the T-region of the nicked Ti plasmid to free the T-strand molecule and produce a Ti plasmid whose T-region is momentarily ss prior to replacement strand synthesis. Since we observe at most one T-strand molecule per AS-induced cell, T-strand production must be tightly regulated.

The sequences internal to the T-DNA of the wild-type Ti plasmid encode genes whose expression in the transformed plant cell result in the tumorous phenotype, crown gall¹, and when the orientation of the right border on the Ti plasmid is flipped (by *in vitro* manipulation), phenotypically transformed plant cells are not obtained^{15,16}. The model of Fig. 6A explains this functional polarity of the right T-DNA border; that is, the T-strand corresponds to the bottom strand of the T-region and must be generated in a right-to-left (5' to 3') direction. Thus, if the orientation of the right border is reversed in the Ti plasmid, ss molecules will be generated away from (rightward of) the T-DNA tumour genes. This model may also explain how the ds T-DNA circle molecules, recovered in *E. coli* transformed with AS-induced DNA, are generated at a low frequency in response to AS induction. Since nicked DNA stimulates recombination events²⁸, the S_1 border sites could promote pairing and recombination

between the T-DNA border repeats of the AS-induced Ti plasmid (Fig. 6B).

Assuming that the T-strand is the transfer intermediate, we can speculate on the mechanism of its transfer. This mechanism might share features with known bacterial processes which mediate the transfer of specific DNA molecules between bacteria, such as phage infection or conjugation. An important distinction between these two processes is that only phage infection involves the synthesis of many copies of the molecule destined for transfer. Since the T-strand is present at about one copy per AS-induced cell, it is unlikely that it is transferred to the plant cell as an infectious phage particle.

In bacterial conjugation, one strand of a ds donor molecule is transferred as a linear ss molecule from the donor to recipient cell²⁹. This process is initiated by nicking one strand of the donor molecule at a specific site (designated *oriT*, origin of transfer)³⁰, and the strand destined for transfer is unwound in a 5'→3' direction. Concomitant to unwinding, the unwound strand is mobilized to the recipient cell, and this transfer is accompanied by DNA synthesis on the donor molecule to replace the mobilized strand. The T-DNA homologous molecules that we have described correspond to the structures which would be predicted to occur if T-DNA transfer occurs through a conjugative mechanism. The S_1 border sites are analogous to nicked *oriT* sites, the T-strand is analogous to the linear ss DNA molecule transferred during bacterial conjugation, and the internal T-region structures are analogous to the replacement strand synthesis intermediates of the donor molecule. Furthermore, bacterial conjugation requires direct contact between donor and recipient cells³¹, and the same requirement is observed for the T-DNA transfer process³².

If the T-strand is transferred by conjugation to the plant cell, it is still not known how it ultimately finds its way into the plant cell nucleus and becomes integrated into the nuclear genome. While the deproteinized T-strand that we have described is a naked linear molecule, presumably it is transferred as a complex that carries proteins which in part mediate the post-transfer events. These proteins, as well as the proteins involved in the generation and transfer of the T-strand, are probably encoded by the plant-inducible Ti-plasmid *vir* loci¹⁰.

We thank Marc Van Montagu for his encouragement and suggestions for the exonuclease studies, Kan Wang for helpful discussion, Jeff Schell for support, and Martine De Cock, Karel Spruyt and Albert Verstraete for preparation of this manuscript and figures. This work was supported by grants from the 'ASLK-Kankerfonds', the 'Fonds voor Geneeskundig Wetenschappelijk Onderzoek' (FGWO 3.001.82), and the Services of the Prime Minister (OOA 12.0561.84) (to Jeff Schell and Marc Van Montagu).

Note added in proof: The nondenatured transfer assay for single-stranded DNAs has also been recently described by Reile, H. T., Michel, B. and Ehrlich, S. D. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2541-2545 (1986).

Received 24 April; accepted 19 June 1986.

- Gheysen, G., Dhaese, P., Van Montagu, M. & Schell, J. in *Advances in Plant Gene Research* Vol. 2 (eds Hohn, B. & Dennis, E. S.) 11-47 (Springer, Vienna, 1985).
- Zambryski, P., Depicker, A., Kruger, K. & Goodman, H. *J. molec. appl. Genet.* **1**, 361-370 (1982).
- Yadav, N. S., Vanderleyden, J., Bennett, D. R., Barnes, W. M. & Chilton, M.-D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6322-6326 (1982).
- Simpson, R. B. *et al.* *Cell* **29**, 1005-1014 (1982).
- Holsters, M. *et al.* *Molec. gen. Genet.* **190**, 35-41 (1983).
- Zambryski, P. *et al.* *EMBO J.* **2**, 2143-2150 (1983).
- de Framond, A. J., Barton, K. A. & Chilton, M.-D. *Bio/technology* **1**, 262-269 (1983).
- Bevan, M. *Nucleic Acids Res.* **12**, 8177-8721 (1984).
- An, G., Watson, B. D., Stachel, S., Gordon, M. P. & Nester, E. W. *EMBO J.* **4**, 277-284 (1985).
- Stachel, S. E. & Nester, E. W. *EMBO J.* **5**, 1445-1454 (1986).
- Douglas, C. J., Staneloni, R. J., Rubin, R. A. & Nester, E. W. *J. Bact.* **161**, 850-860 (1985).
- Stachel, S. E., Nester, E. W. & Zambryski, P. C. *Proc. natn. Acad. Sci. U.S.A.* **83**, 379-383 (1986).
- Stachel, S. E. & Zambryski, P. *Cell* (in the press).
- Stachel, S. E., Messens, E., Van Montagu, M. & Zambryski, P. *Nature* **318**, 624-629 (1985).
- Wang, K., Herrera-Estrella, L., Van Montagu, M. & Zambryski, P. *Cell* **38**, 455-462 (1984).
- Peralta, E. G. & Ream, L. W. *Proc. natn. Acad. Sci. U.S.A.* **82**, 5112-5116 (1985).
- Joos, H. *et al.* *Cell* **32**, 1057-1067 (1983).

18. Shaw, C. H., Watson, M. D., Carter, G. H. & Shaw, C. H. *Nucleic Acids Res.* **12**, 6031-6041 (1984).
19. Koukoliková-Nicola, Z. *et al.* *Nature* **313**, 191-196 (1985).
20. Southern, E. M. *J. molec. Biol.* **98**, 503-518 (1975).
21. Bourgeois, S. *Meih. Enzym.* **21**, 491-500 (1971).
22. Vogt, V. *Eur. J. Biochem.* **33**, 192-200 (1973).
23. Nishigaki, K., Kaneko, Y., Wakuda, H., Husimi, Y. & Tanaka, T. *Nucleic Acids Res.* **13**, 5747-5760 (1985).
24. Huang, W. M. & Lehman, I. R. *J. biol. Chem.* **247**, 3139-3146 (1972).
25. Chase, J. W. & Richardson, C. C. *J. biol. Chem.* **249**, 4545-4561 (1974).
26. Lilley, D. M. J. *Nucleic Acids Res.* **9**, 1271-1289 (1981).
27. Panayotatos, N. & Wells, R. D. *Nature* **289**, 466-470 (1981).
28. Dressler, D. & Potter, H. A. *Rev. Biochem.* **51**, 727-761 (1982).
29. Willetts, N. & Wilkins, B. *Microbiol. Rev.* **48**, 24-41 (1984).
30. Thompson, R., Taylor, L., Kelly, K., Everett, R. & Willetts, N. *EMBO J.* **3**, 1175-1180 (1984).
31. Manning, P. A. & Achtman, M. in *Bacterial Outer Membranes* (ed. Inouye, M.) 409-447 (Wiley, New York, 1979).
32. Douglas, C. J., Halperin, W. & Nester, E. W. *J. Bact.* **152**, 1265-1275 (1982).
33. Dhaese, P., De Greve, H., Decraemer, H., Schell, J. & Van Montagu, M. *Nucleic Acids Res.* **7**, 1837-1849 (1979).
34. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning, a Laboratory Manual* (Cold Spring Harbor Laboratories, New York, 1982).

LETTERS TO NATURE

Binary pulsar with a very small mass function

R. J. Dewey*, C. M. Maguire†, L. A. Rawley†, G. H. Stokes† & J. H. Taylor†

* Center for Radiophysics and Space Research, Cornell University, Ithaca, New York, 14853, USA

† Joseph Henry Laboratories and Physics Department, Princeton University, Princeton, New Jersey 08544, USA

Much of the interesting physics concerning neutron stars and their evolution depends for its experimental foundation on observations of the rotation rates of pulsars. To continue recent efforts of our group in this area¹⁻³, we began a series of pulse-arrival-time observations of ~70 pulsars in January 1985. Most of the pulsars in this study were discovered in the Princeton/NRAO pulsar survey of the preceding two years^{4,5}. Soon after we began these observations it became clear that PSR2303+46 was a binary pulsar²; it is now evident that PSR1831-00 is also a member of a binary system, the seventh such radio pulsar known. It moves in an orbit with a period of 1.81 days, a small eccentricity, and an unusually small mass function of $0.00012 M_{\odot}$ (where $M_{\odot}$ is the mass of the Sun). With a period $P = 0.521$ s and period derivative $\dot{P} \approx 10^{-17}$ s s⁻¹, PSR1831-00, like the other known binary pulsars, has a relatively weak magnetic field. We discuss the features of this system that provide clues to its evolutionary history and outline possible models for its formation.

Most of our observations were made with the 92-m transit telescope at Green Bank, West Virginia, at a frequency of 390 MHz. The data acquisition system has been described by Stokes *et al.*⁵. Briefly, a dual-channel parametric up-converter amplifies two orthogonal linear polarizations and provides a system noise temperature of 50 K at high galactic latitudes. In the direction of PSR1831-00 (galactic coordinates $l = 30.8^{\circ}$, $b = 3.7^{\circ}$) the system temperature is 180 K, equivalent to a flux density of ~150 Jy. In each polarization an 8-MHz pass-band is divided into 32 sub-channels, each 250 kHz wide; the signals are detected, summed in a de-disperser, and then integrated for two minutes in a signal averager synchronized to the apparent pulsar period. For each pulsar, the resulting profiles are cross-correlated with a standard profile for that pulsar. This procedure determines phase offsets which, when added to reference times near the centre of the integrations, yield effective pulse arrival times.

For PSR1831-00, one to ten arrival times were obtained in this way on each of 19 days in January, February, April, July and November 1985, and February 1986. The number of observations obtainable on a single day is limited by the small hour-angle range through which the 92-m telescope can track (~20 min at the celestial equator), and this in turn limits the accuracy with which one can measure a pulsar's apparent period on a given day. Nevertheless, for virtually all non-binary pulsars it is possible to fit data obtained over two or three days to a single barycentric period and to count pulses unambiguously

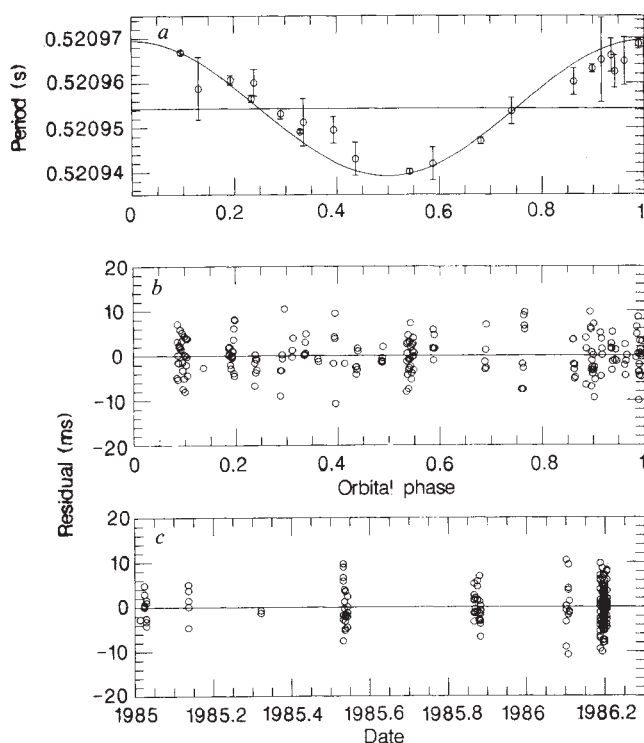


Fig. 1 Pulsar period as a function of orbital phase (a), and arrival-time residuals as function of both orbital phase (b) and date (c). The sinusoid superimposed on the measured periods in a corresponds to the orbital parameters listed in Table 1.

from day to day. The data for PSR1831-00 were not amenable to such treatment; it was rarely possible to fit data taken one or two days apart to a single period. However, the day-to-day period changes were small (a few parts in 10^5) and only marginally significant, making the binary nature of this pulsar hard to recognize.

These difficulties were compounded by the weakness of the pulsar's signal (2-5 mJy) and its variability. We did not detect the pulsar in a number of our attempts to observe it, and there is some evidence that it has become weaker in the past year. The variability may be due to refractive interstellar scintillation. Our failures to detect the pulsar show no correlation with orbital phase, so there is no evidence that eclipses are involved.

Redoubled observational efforts in February 1986 showed that the data were consistent with a nearly circular orbit having a 1.81-day period and a maximum radial velocity of ± 8.7 km s⁻¹. We were still not able, however, to connect pulsar phases unambiguously between observing days. Consequently, during March 1986 we observed PSR1831-00 at 430 MHz with the Arecibo 305-m telescope. The data acquisition scheme is described in ref. 6. The pulsar is close to the southern declination limit of the Arecibo antenna and can therefore be observed for

only 50 min each day; however, the improvement in tracking time over that available at Green Bank allowed us to connect pulsar phases between observations made on 10, 12, 13, 14 and 16 March. An orbital solution obtained using the method described by Manchester and Taylor⁷ was good enough to allow unambiguous phase connection to all of the earlier Green Bank data. The resulting global solution for pulsar and orbital parameters is presented in Table 1, and the orbital velocity curve

Table 1 Parameters of the PSR1831-00 system

Right ascension (1950.0)	$\alpha = 18^{\text{h}} 31^{\text{m}} 43.23 \pm 0.08^{\text{s}}$
Declination (1950.0)	$\delta = -00^{\circ} 13' 13.5 \pm 1.8''$
Pulsar period	$P = 0.5209543084(3)\text{s}$
Period derivative	$\dot{P} = (0.2 \pm 1.6) \times 10^{-17} \text{s s}^{-1}$
Dispersion measure	$DM = 94 \pm 6 \text{ cm}^{-3} \text{ pc}$
Projected semi-major axis	$a_1 \sin i = 0.7227 \pm 0.0010 \text{ light s}$
Eccentricity	$e < 0.005$
Orbital period	$P_b = 156478.9 \pm 0.4 \text{ s}$
Time of periastron	$T_o = 2,446,460.3411 \pm 0.0011 \text{ JD}$
Assumed longitude of periastron	$\omega = 0.0^{\circ}$

and post-fit arrival-time residuals are shown in Fig. 1. We note that our measured period derivative, although effectively an upper limit, already places PSR1831-00 among the dozen pulsars with the smallest slow-down rates, as can be seen in Fig. 2. A small rate of spin-up is also consistent with our data—and would be most interesting if confirmed—but is not yet required. At present only an upper limit is available for the orbital eccentricity. If it is as small as the eccentricities of the PSRs 0655 + 64, 1855 + 09 and 1953 + 29, it will be impossible to measure because the ratio $(a_1 \sin i)/P$ for PSR1831-00 is so small. (a_1 is the orbital semi-major axis; i is the angle between the plane of the orbit and the plane of the sky.)

The orbital period and semi-major axis determine the mass function:

$$f(m_1, m_2) = \frac{(m_2 \sin i)^3}{(m_1 + m_2)^2} = \frac{4\pi^2 (a_1 \sin i)^3}{G P_b^2} = 0.000123 M_{\odot}$$

where m_1 and m_2 are the masses of the pulsar and companion, respectively. As shown in Fig. 3, this constraint implies a very small companion mass, $\sim 0.05\text{--}0.15 M_{\odot}$, unless $\cos i$ is implausibly close to 1.

As can be seen in Table 2, the 1831-00 system is similar in many respects to the other known binary pulsars. Like the other binaries, the pulsar has an unusually small spin-down rate and an unusually small inferred magnetic field, $B = 3.2 \times 10^{19} (P\dot{P})^{1/2} < 8 \times 10^{10} \text{ G}$. All seven binary pulsars lie between the 'spin-up line' and the 'death line'^{3,8} in the $P\text{--}\dot{P}$ diagram, as shown in Fig. 2. Taylor and Stinebring⁹ note that the known binary pulsars and the millisecond pulsar 1937 + 21 show a correlation between magnetic field and distance $|z|$ from the galactic plane, a pattern into which PSR1831-00 fits nicely (see Table 2). Although this relation may be partly due to selection effects—searches for millisecond pulsars (expected to have small magnetic fields) have tended to concentrate at low latitudes—it may also provide information about how binary

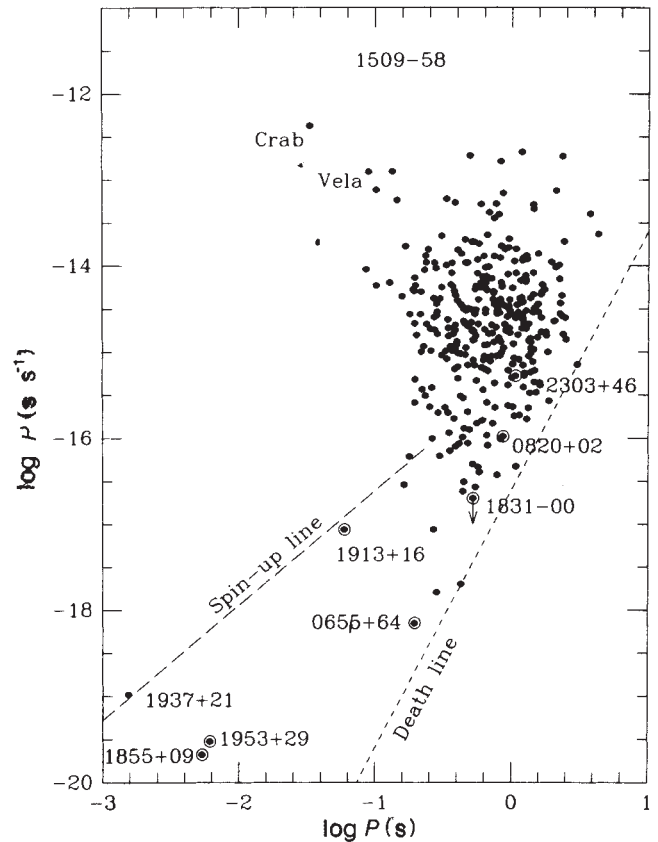


Fig. 2 Distribution of periods and period derivatives for 353 pulsars. The seven known binary pulsars, indicated by circles around the dots, have unusually small period derivatives and hence relatively weak magnetic fields.

and millisecond pulsars are formed, and may be related to the correlation between magnetic field and velocity first noted by Anderson and Lyne¹⁰. This correlation does not, however, follow simply from models of binary pulsar formation such as those outlined by van den Heuvel⁸.

What makes the 1831-00 system atypical, and its evolutionary history somewhat difficult to understand, is its relatively tight orbit (component separation $\sim 6 R_{\odot}$, where $R_{\odot}$ is the radius of the Sun) and extremely small mass function. Models for the formation of binary pulsars⁸ suggest that during evolution a low-mass companion will spiral away from a neutron-star primary, leaving a final system with a wide orbit similar to those of PSRs 1953 + 29 and 0820 + 02. Nevertheless, a number of features of the evolution of the 1831-00 system may be inferred from its orbital parameters and the pulsar's period and spin-down rate.

The pulsar's small magnetic field suggests that it is relatively old ($\geq 3 \times 10^7$ yr, assuming a magnetic field decay time of 10^7 yr and a magnetic field of 10^{12} G at birth¹¹). As the system presently

Table 2 Parameters of binary and millisecond pulsars

PSR	P (ms)	$\log \dot{P}$	$\log B$ (G)	$ z $ (pc)	$a_1 \sin i$ ($R_{\odot}$)	P_b (days)	e	$f(m_1, m_2)$ ($M_{\odot}$)	Likely m_2 ($M_{\odot}$)
1937 + 21	1.6	-19.0	8.6	20	—	—	—	—	—
1855 + 09	5.4	-19.7	8.5	20	4.0	12.33	0.00002	0.0052	0.2-0.4
1953 + 29	6.1	-19.5	8.6	20	13.6	117.35	0.0003	0.0027	0.2-0.4
0655 + 64	195.6	-18.2	10.0	120	1.8	1.03	<0.00005	0.0712	0.7-1.3
1913 + 16	59.0	-17.1	10.3	190	1.0	0.32	0.6171	0.1322	1.4
1831 - 00	520.9	< -17.0	<10.9	190	0.3	1.81	<0.005	0.00012	0.06-0.13
0820 + 02	864.9	-16.0	11.5	280	70.0	1,232.40	0.0119	0.0030	0.2-0.4
2303 + 46	1,066.4	-15.4	11.8	480	14.1	12.34	0.6584	0.2463	1.2-2.5

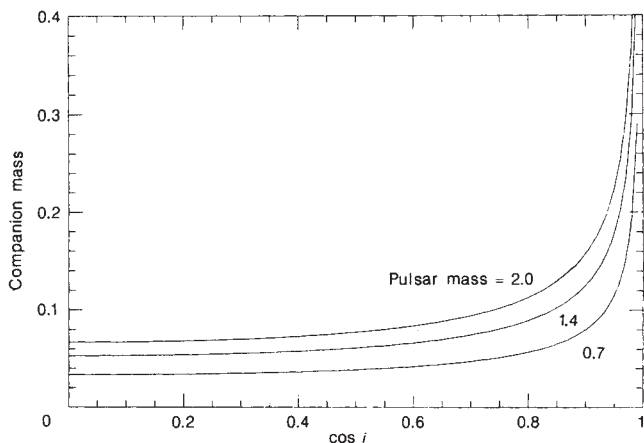


Fig. 3 Mass of the companion of PSR1831-00, plotted as a function of $\cos i$ and assuming three different values for the mass of the pulsar. All masses are in units of $M_{\odot}$.

lies ~ 200 pc above the galactic plane, such an age implies a velocity of $\leq 10 \text{ km s}^{-1}$, substantially smaller than that of the average pulsar. A shorter magnetic field decay time or smaller initial magnetic field would allow a larger velocity. The position of the pulsar in the $P-\dot{P}$ diagram (Fig. 3) allows, but does not require, that the pulsar has been spun up by accretion from its companion.

The system's circular orbit is compatible with the companion having undergone mass loss after the pulsar was created. If this is the case, the pulsar may have been formed by the accretion-induced collapse of a white dwarf, which would explain the system's low apparent velocity. However, it is also possible that no mass transfer has occurred since the formation of the neutron star in a supernova explosion, and that tidal forces have circularized the orbit. If the companion has a substantial convective envelope, such forces may be able to do this quite effectively¹².

For $\cos i \leq 0.95$ the companion mass is too small for the companion to be the white-dwarf or helium-core remnant of a star that has evolved off the main sequence in the age of the universe. If the companion is the remnant of an evolved star which underwent mass loss, that mass loss must have occurred before the end of the star's hydrogen burning, and before a sizeable compact core could develop.

This evidence points to two alternative evolutionary models, one involving mass transfer after (or during) the formation of the neutron star, and one in which no mass transfer has occurred. In both models the neutron star is formed from the primary of the original main-sequence system. In the former case, because the companion must have overflowed its Roche lobe before leaving the main sequence, the orbit after the collapse of the primary must have been quite small. This suggests that the progenitor system passed through a contact phase during the primary's evolution. During this phase the secondary may have accreted a significant amount of matter, but this is not required. Once the secondary overflowed its Roche lobe, the system may have appeared as a cataclysmic variable or low-mass X-ray binary and during this stage a white dwarf might be pushed over the Chandrasekhar limit. The compact primary's motion must have acted to eject the envelope of the secondary, a process which appears to be possible, given suitable conditions¹³. If the pulsar was formed by accretion-induced collapse, the most likely pre-neutron-star candidates are an O-Ne-Mg white dwarf whose progenitor was a fairly massive ($\geq 8 M_{\odot}$) star, or a very old CO white dwarf¹⁴.

In the alternative model, the companion is a normal low-mass star which has evolved very little since the formation of the pulsar in a supernova explosion. In order for the system to have remained bound, the primary must have had a mass of $\sim 3 M_{\odot}$ when it exploded, unless very limiting assumptions are made

about asymmetries in the supernova explosion. This again requires a contact phase during the evolution of the primary, in this case, one in which the companion accreted very little matter. The explosion of any helium star massive enough to have formed a neutron star would have induced a substantial (≥ 0.5) eccentricity in the system's orbit, so this model is viable only if tidal forces could have circularized the orbit within the lifetime of the pulsar.

We are grateful to T. H. Hankins, D. R. Stinebring and K. Turner for allowing us to use the Arecibo telescope during part of their scheduled time, and to J. M. Cordes and A. Wolszczan for help with some of the observations. We thank the staffs of the National Radio Astronomy Observatory and the National Astronomy and Ionosphere Center for assistance with the observations and NSF for financial support. R.J.D. thanks E. E. Salpeter for informative discussions on the evolution of low-mass stars.

Received 18 April; accepted 13 June 1986.

1. Davis, M. M., Taylor, J. H., Weisberg, J. M. & Backer, D. C. *Nature* **315**, 547-550 (1985).
2. Stokes, G. H., Taylor, J. H. & Dewey, R. J. *Astrophys. J.* **294**, L21-L24 (1985).
3. Rawley, L. A., Taylor, J. H. & Davis, M. M. *Nature* **319**, 383-384 (1986).
4. Dewey, R. J., Taylor, J. H., Weisberg, J. M. & Stokes, G. H. *Astrophys. J.* **294**, L25-L29 (1985).
5. Stokes, G. H., Taylor, J. H., Weisberg, J. M. & Dewey, R. J. *Nature* **317**, 787-788 (1985).
6. Segelstein, D. J., Rawley, L. A., Stinebring, D. R., Fruchter, A. S. & Taylor, J. H. *Nature* (1986).
7. Manchester, R. N. & Taylor, J. H. *Pulsars* (Freeman, San Francisco, 1977).
8. van den Heuvel, E. P. J. *J. Astrophys. Astr.* **5**, 209-233 (1984).
9. Taylor, J. H. & Stinebring, D. R. *A. Rev. Astr. Astrophys.* (in the press).
10. Anderson, B. & Lyne, A. G. *Nature* **303**, 597-599 (1983).
11. Lyne, A. G., Manchester, R. N. & Taylor, J. H. *Mon. Not. R. astr. Soc.* **213**, 613-639 (1985).
12. Lecar, M., Wheeler, J. C. & McKee, C. F. *Astrophys. J.* **205**, 556-562 (1976).
13. Tamm, R. E., Bodenheimer, P. & Ostriker, J. P. *Astrophys. J.* **222**, 269-280 (1978).
14. Trimble, V. J. *Astrophys. Astr.* **5**, 389-402 (1984).

New millisecond pulsar in a binary system

D. J. Segelstein, L. A. Rawley, D. R. Stinebring, A. S. Fruchter & J. H. Taylor

Joseph Henry Laboratories and Physics Department, Princeton University, Princeton, New Jersey 08544, USA

Recent observations^{1,2} at the Arecibo Observatory have resulted in the discovery of PSR1855+09, a pulsar with period $P = 5.362$ ms, moving in a nearly circular orbit of period 12.3 days. The pulsar is only the third one known with $P < 10$ ms, and the sixth known radio pulsar in a binary system. (Discovery of a seventh binary pulsar is announced in an accompanying paper³.) Three of the seven binaries are among the fastest five of more than 400 pulsars—a fact that provides strong support for the conclusion that fast pulsars are 'recycled' neutron stars, spun up during a phase of mass accretion from an evolving companion star. The pulsar has a small dispersion measure ($13.3 \text{ cm}^{-3} \text{ pc}$), suggesting a distance of only ~ 350 pc. The proximity of this pulsar and its location within 25 pc of the galactic plane argue that millisecond pulsars form a significant fraction ($\sim 10\%$) of the pulsar population, leaving many detectable ones undiscovered^{4,5}. Its signal is strong enough to permit pulse-arrival-time measurements with single-day uncertainties of $< 3 \mu\text{s}$. Timing observations already suggest that PSR1855+09, like the 1.5-ms pulsar PSR1937+21, will prove to be a natural clock of extremely high stability⁶. The existence of a second pulsar with extremely small timing uncertainties will greatly aid the search for background gravitational waves using millisecond pulsars as detectors.

Observations leading to the discovery of PSR1855+09 were part of a survey specifically designed to detect fast pulsars¹. The Arecibo 305-m telescope was used at a frequency of 430 MHz, where a 30-m line feed provides a sensitivity of up to 19 K Jy^{-1}

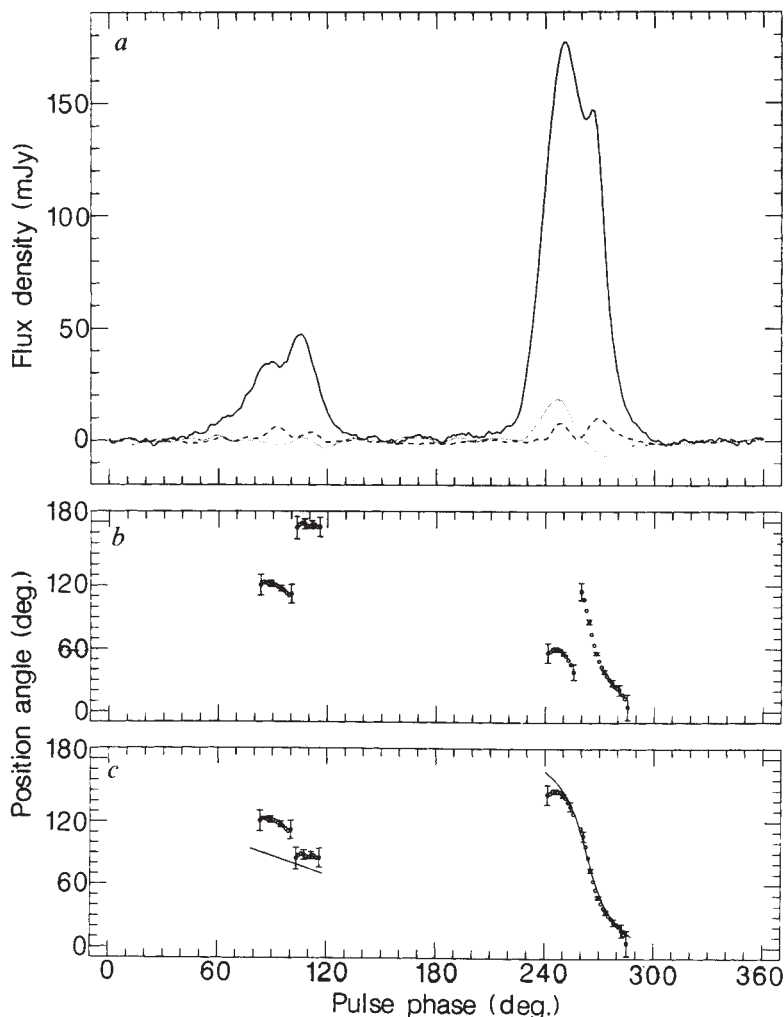


Fig. 1 Total intensity and polarization profiles for PSR1855+09. *a*, Total intensity (solid line), linear polarization (dashed line) and circular polarization (dotted line) for a synchronous average of $\sim 560,000$ pulses. There are 256 samples across the pulse ($20.9 \mu\text{s}$ per sample); dispersion smearing was $40 \mu\text{s}$ across a 1 MHz bandwidth and the post-detection time constant was $50 \mu\text{s}$. The total intensity data are shown at full resolution, whereas the linear and circular polarization data have been boxcar-smoothed over four time samples. *b*, Position angle of the linear polarization, plotted wherever the linear polarization was >3.0 times the r.m.s. (root mean square) noise. Representative error bars are ± 1.0 the r.m.s. error in the position angle. *c*, Position angle, with two segments shifted by 90° , plotted together with a representative fit to a Radhakrishnan and Cooke¹⁵ model ($\alpha = 30^\circ$, $\beta = 25.5^\circ$, $\theta_0 = 91^\circ$ and $\phi_0 = 263^\circ$).

in each sense of circular polarization and a half-power beam-width of $10'$. At zenith angle 10° and galactic background temperature 150 K, parameters typical of the region of sky surveyed, the flux density equivalent to the total system noise was ~ 17 Jy.

During the survey we observed a grid of 14,452 beam areas, with $9'$ spacing between points, for 39 s each. A filter-bank spectrometer provided 16 adjacent channels of 60 kHz width for each circular polarization. Detected power from the two polarizations was summed for each channel, and the sixteen resulting signals were smoothed with a time constant of $400 \mu\text{s}$, sampled every $300 \mu\text{s}$, digitized to 3-bit precision, and stored on magnetic tape for later analysis. Further details of the searching procedure are contained in ref. 1. The survey observation in the direction of PSR1855+09 was made in October 1984, and confirming observations followed in July and November 1985. By mid-December it was clear that a new millisecond pulsar had been found, and that its period was changing significantly on a timescale of several days. The pulsar's mean flux density at 430 MHz is ~ 10 mJy.

Starting on 20 December 1985, we have made timing measurements of the new pulsar with equipment and techniques similar to those used to observe PSR1953+29⁷. The Arecibo telescope was used for these observations, usually at 1,400 MHz, with a $2 \times 32 \times 250$ kHz filter bank. Because interstellar scintillation causes the signal strength to vary widely (with a decorrelation bandwidth of ~ 8 MHz, a timescale of ~ 40 min, and a peak-to-mean flux density ratio approaching 10), we choose an exact observing frequency, somewhere in the range 1,360–1,440 MHz, at the time of each observation. A 32-channel signal averager samples the total-power signals 256 times per pulsar period and records 32 integrated pulse profiles every two minutes. Up to

45 such observations can be made on a given day. Cross-correlation of the observed profiles with a standard profile produces a set of phase delays which are corrected for dispersion, averaged and added to a reference time near the midpoint of the integration to yield an effective pulse arrival time. Based on the variance of dispersion-corrected phase delays, the uncertainties in arrival times for 2-min integrations are between 2 and $15 \mu\text{s}$, depending on signal strength. The total-intensity pulse profile is shown in Fig. 1*a*, together with the linear and circular polarization.

A 15-min observation made with the Very Large Array (VLA) on 28 February 1986, using two 50-MHz bandwidths centred at 1,464.9 MHz and 1,514.9 MHz, found an average pulsar flux density of 1.5 mJy, and provided a position measurement in the FK4 coordinate system of $\alpha(1950) = 18 \text{ h } 55 \text{ min } 13.681 \pm 0.009 \text{ s}$, $\delta(1950) = +09^\circ 39' 12.80 \pm 0.15''$. Because the product of the decorrelation bandwidth and the peak-to-mean flux ratio is as large as our total bandwidth, the measured flux over a 15-min period is subject to large statistical fluctuations. Our estimated flux density, therefore, is accurate only to about a factor of two. For these observations we used the radio source 1821+07 as a phase calibrator; its position was assumed to be $\alpha(1950) = 18 \text{ h } 21 \text{ min } 41.655 \text{ s}$, $\delta(1950) = +10^\circ 42' 43.90''$.

The timing data were analysed using a procedure similar to that described by Rawley *et al.*⁷. We first determined the best-fit Solar System barycentric period for each available 20-min interval of data. These periods were used to carry out a least-squares solution for the pulsar period in its own rest frame, P , the binary orbital period, P_b , the projected semi-major axis, $a_1 \sin i$ (where i is the angle between the plane of the orbit and the plane of the sky), and a time of passage through the

ascending node. The sinusoidal velocity curve determined from these parameters is superimposed on a plot of the barycentric period measurements at the top of Fig. 2.

Unambiguous pulse numbers could then be assigned to the 1,073 pulse arrival times obtained between 20 December 1985 and 24 May 1986. A least-squares fit yielded the pulsar parameters and orbital elements listed in Table 1. Figure 2 shows

Table 1 Parameters of the PSR1855+09 system

Right ascension (1950.0)	$\alpha = 18^{\text{h}} 55^{\text{m}} 13.6834 \pm 0.0006^{\text{s}}$
Declination (1950.0)	$\delta = +09^{\circ} 39' 13.278 \pm 0.009''$
Pulsar period	$P = 5,362,100,452.39 \pm 0.03^{\text{ps}}$
Period derivative	$\dot{P} = (2.1 \pm 0.5) \times 10^{-20} \text{ s s}^{-1}$
Dispersion measure	$DM = 13.2943 \pm 0.0002 \text{ cm}^{-3} \text{ pc}$
Projected semi-major axis	$a_1 \sin i = 9.2307850 \pm 0.0000004 \text{ light s}$
Eccentricity	$e = (21.27 \pm 0.08) \times 10^{-6}$
Orbital period	$P_b = 1,065,067.593 \pm 0.004^{\text{s}}$
Longitude of periastron	$\omega = 277.10210 \pm 0.33^{\circ}$
Time of periastron	$T_0 = 2,446,433.3026117 \pm 0.0114 \text{ JD}$

post-fit residuals for this solution, plotted as a function of both data and orbit phase. The timing position shown in Table 1 is based on the coordinate system of the PEP740-R ephemeris (J. F. Chandler *et al.*, personal communication). Positions obtained from this ephemeris must be rotated by $\sim 0.4''$ before they can be compared with FK4 coordinates^{8,9}. After this is done, the discrepancy between the VLA and timing positions is about two standard deviations—probably not significant.

The very small period derivative ($\dot{P}$) observed for PSR1855+09 implies a surface magnetic field strength of $\sim 3 \times 10^8 \text{ G}$ —very close to the values obtained for the other two millisecond radio pulsars^{6,7}, and consistent with binary spin-up models for their evolution (refs 10, 11 and references therein). A small $\dot{P}$ also suggests that the rotational stability of this pulsar, like that of PSR1937+21, is probably very high⁶. This conjecture is supported by the fact that after averaging our timing data to give a single equivalent arrival time for each of the 21 observing days, we obtain a post-fit weighted r.m.s. residual of only 2.6 μs . Future comparison of timing residuals of PSR1855+09 with those of PSR1937+21 may be able to distinguish effects external to the Solar System (such as a cosmic background of gravitational radiation⁶) from inaccuracies in the Earth's ephemeris or in the realization of atomic time.

Because of the extremely small eccentricity of the orbit of PSR1855+09, relativistic effects will not be readily detectable in its timing data. General relativistic advance of periastron should occur at a rate of $\sim 15'' \text{ yr}^{-1}$, but with present accuracies it will require a century or more to detect. Some further conclusions concerning statistics of the population of binary radio pulsars are discussed in ref. 3.

On 7 and 18 January 1986 we made phase-resolved polarization observations of the new pulsar at 1,400 MHz, using a procedure similar to that described by Stinebring¹². The mean polarization of this pulsar at 1,400 MHz reaches a maximum of $\sim 10\%$ circular (near the peak of the main pulse) and 10% linear (on the trailing edge). The circular polarization changes sign at the centre of the main pulse and the linear polarization goes to zero between the components of the main pulse and interpulse, indicating the presence of orthogonally polarized radiation^{13,14}.

The position-angle swing of the linear polarization of many pulsars is consistent with a simple model first put forward by Radhakrishnan and Cooke¹⁵, in which the position angle at a particular rotational phase is determined by the projection of the local magnetic field direction onto a plane perpendicular to the line of sight. Assuming a purely dipolar magnetic field structure, the position angle χ is given by¹⁶

$$\tan(\chi - \chi_0) = \frac{\sin \alpha \sin(\phi - \phi_0)}{\cos \alpha \sin \beta - \sin \alpha \cos \beta \cos(\theta - \theta_0)}$$

where α is the angle between the rotational axis and the magnetic dipole axis, β is the angle between the rotational axis and the

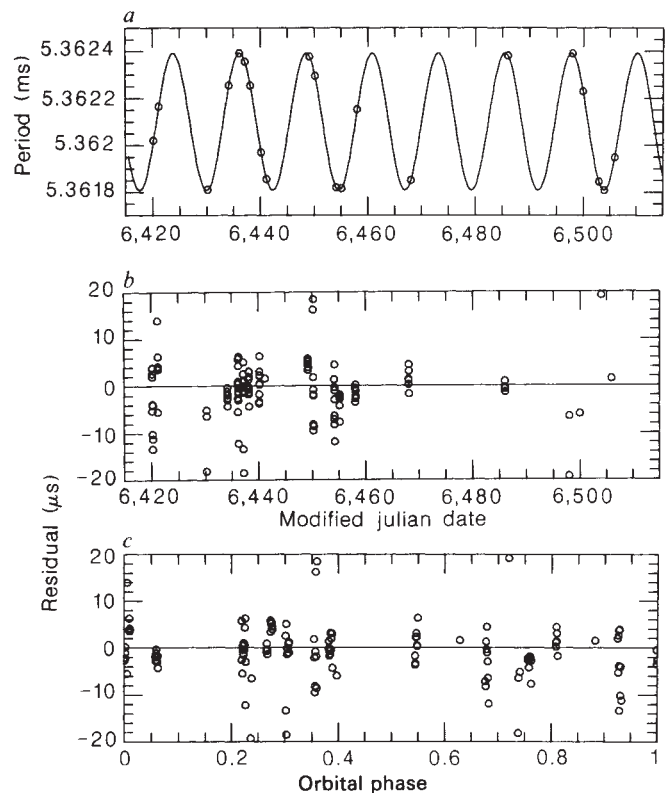


Fig. 2 Summary of PSR1855+09 timing data, 20 December 1985 to 16 March 1986. *a*, Sinusoidal fit to pulse period measurements (circles) on 21 different days. For the sake of clarity only one period measurement per day is plotted. *b*, *c*, Residuals from fit (Table 1) to pulse arrival times, plotted against time in *b* and orbital phase (fraction of an orbit since periastron passage) in *c*. Each circle corresponds to 10 min of observing time. No attempt is made to distinguish among data of widely varying quality.

line of sight, and ϕ is the azimuth of the line of sight relative to the projection of the magnetic axis onto the equatorial plane. The angles ϕ_0 and χ_0 are physically uninteresting, as the origin of pulse phase and position angle are arbitrary.

A fit of this equation to the position-angle data is shown in Fig. 1c, where the position angle has been shifted by 90° on the trailing half of the interpulse and the leading half of the main pulse, assuming that this radiation is in an orthogonal polarization state. A least-squares fit for α and β was performed after fixing $\phi_0 = 263^\circ$ and $\chi_0 = 91^\circ$ because of the obvious symmetry point at the centre of the main pulse. Over the range $10^\circ \leq \alpha \leq 90^\circ$ we find good fits to the main-pulse position-angle sweep for an intercept angle $\alpha - \beta \approx q\alpha$, where q varies smoothly from 0.2 at $\alpha = 10^\circ$ to 0.1 at $\alpha = 90^\circ$. Although the fit shown is not very good in the interpulse region, the position-angle swing does match the slope of the interpulse data fairly well. Any fit with $\beta > \alpha$ can be ruled out because the observed position-angle gradient has the same sign across the interpulse and the main pulse.

The position angle data are consistent with an interpulse produced at either the same magnetic pole as the main pulse or at the opposite magnetic pole¹⁷. The symmetry of the main-pulse and interpulse morphology, as well as the indication that corresponding components of the main pulse and interpulse are in the same polarization mode, lead us to favour a one-pole model similar to that proposed by Narayan and Vivekanand¹⁸ for PSR0950+08. Their model, in which α is relatively small and we cut across both edges of a hollow cone twice, is in good qualitative agreement with the polarization and pulse morphology data for this pulsar. Such a model is made more likely by the evidence that pulsar beams are elongated in the direction

of latitude and that this elongation is more marked for shorter-period pulsars¹⁹. The deviation of the interpulse position angle from the simple model considered here may arise from a departure from dipolar magnetic field geometry, magnetospheric propagation effects²⁰, or interpulse emission from a much higher altitude than for the main pulse²¹.

The discovery of a third millisecond pulsar so nearby indicates that millisecond pulsars, although difficult to find, are not rare objects in the pulsar population. As their presumed progenitors—the low-mass X-ray binary systems—are rare, millisecond pulsars may be observable for much longer than the 10^8 -yr lifetimes of the X-ray systems. This, in turn, implies that pulsar magnetic fields may not decay below a level of $\sim 5 \times 10^8$ G within a Hubble time, an idea discussed in detail by other authors^{5,11}. Independent evidence that millisecond pulsars are very old and long-lived comes from the detection of cool white dwarf companions for PSR0655+64 and PSR0820+02 by Kulkarni²², and for the millisecond pulsar reported here by Wright and Loh²³.

We thank the staff of the Arecibo Observatory, and in particular A. Wolszczan, for help with the observations and the U.S. National Science Foundation for financial support. We would also like to thank J. H. van Gorkom for providing observing time and technical assistance at the VLA.

Received 18 April; accepted 13 June 1986.

1. Segelstein, D. J. thesis, Princeton Univ. (1986).
2. Segelstein, D. J., Taylor, J. H., Rawley, L. A., Stinebring, D. R.: Wolszczan, A. *IAU Circ. No. 4262* (1986).
3. Dewey, R. J., Maguire, C. M., Rawley, L. A., Stokes, G. H. & Taylor, J. H. *Nature* **322**, 712–714 (1986).
4. Stokes, G. H., Segelstein, D. J., Taylor, J. H. & Dewey, R. J. *Astrophys. J.* (submitted).
5. Bhattacharya, D. & Srinivasan, G. *Curr. Sci.* **55**, 327–330 (1986).
6. Davis, M. M., Taylor, J. H., Weisberg, J. M. & Backer, D. C. *Nature* **315**, 547–550 (1985).
7. Rawley, L. A., Taylor, J. H. & Davis, M. M. *Nature* **319**, 383–384 (1986).
8. Bartel, N. *et al. Astr. J.* **90**, 318–325 (1985).
9. Backer, D. C., Fomalont, E. B., Goss, W. M., Taylor, J. H. & Weisberg, J. M. *Astr. J.* **90**, 2275–2280 (1985).
10. van den Heuvel, E. P. J. *J. Astrophys. Astr.* **5**, 209–233 (1984).
11. van den Heuvel, E. P. J., van Paradijs, J. & Taam, R. E. *Nature* **322**, 153–155 (1986).
12. Stinebring, D. R. *Nature* **302**, 690–692 (1983).
13. Backer, D. C. & Rankin, J. M. *Astrophys. J. Suppl. Ser.* **42**, 143–173 (1980).
14. Stinebring, D. R., Cordes, J. M., Weisberg, J. M., Rankin, J. M. & Boriakoff, V. *Astrophys. J. Suppl. Ser.* **55**, 247–277 (1984).
15. Radhakrishnan, V. & Cooke, D. J. *Astrophys. Lett.* **3**, 225–229 (1969).
16. Manchester, R. N. & Taylor, J. H. *Pulsars*, 218 (Freeman, San Francisco, 1977).
17. Hankins, T. H. & Cordes, J. M. *Astrophys. J.* **249**, 241–253 (1981).
18. Narayan, R. & Vivekanand, M. *Astrophys. J.* **274**, 771–775 (1983).
19. Narayan, R. & Vivekanand, M. *Astr. Astrophys.* **122**, 45–53 (1983).
20. Barnard, J. *Astrophys. J.* (in the press).
21. Gil, J. *Astrophys. J.* **299**, 154–160 (1985).
22. Kulkarni, S. R. *Astrophys. J.* **306**, L85–L89 (1986).
23. Wright, G. A. & Loh, E. D. *Nature* (in the press).

Repulsive regularities of water structure in ices and crystalline hydrates

Hugh F. J. Savage* & John L. Finney†

* Center for Chemical Physics, National Bureau of Standards, Gaithersburg, Maryland 20899, USA

† Department of Crystallography, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK

Hydrogen-bonded structures display wide ranges of the various angles (such as O—H...O) and distances (such as H...O) used to describe their intermolecular geometry. For example, O...O distances are found to vary between ~ 2.5 and 3.2 Å, and the hydrogen bond angles normally vary between 120 and 180° . Although correlations between distances and angles may be found (for example, between H...O distance and O—H...O angle), these geometrical characteristics are generally too 'soft' to be used as stereochemical restraints in, for example, refining the solvent structure in protein crystals. Moreover, although many potential functions exist for

describing water–water interactions, none of them is totally satisfactory in reproducing experimental results, even for pure water¹. We present here the initial results of an analysis of water structures in high-resolution neutron crystal structures² which is much more successful in rationalizing the stereochemistry of water interactions. Rather than considering the structures in terms of weak, orientation-dependent attractions, a concentration on the repulsive interactions leads to a set of very much stronger stereochemical constraints, which not only rationalize the structures but appear largely to control the orientational correlations in aqueous systems. Looked at this way, water network structures in crystal hydrates and, we suspect, in the liquid itself, become for the first time comprehensible. The approach provides a much firmer base from which to build realistic potential functions to model and simulate solvent structures.

A significant number of atomic-resolution neutron diffraction studies are now available in which the hydrogen (deuterium) atoms can also be located. Figure 1a shows a well-known plot relating to small hydrate crystals³; we see that as the H...O distance of hydrogen bond increases, the mean O—H...O bond angle tends to decrease. However, there is a large scatter in the H-bond geometries about the regression line that can be drawn: some short H-bonds are very bent, while some long ones tend to be very straight. A similar spread of values appears (Fig. 1b) when we restrict ourselves to the H-bond geometries of water molecules in a variety of crystal hydrates (containing between 1 and 18 water molecules per asymmetric unit) and ice polymorphs analysed by high-resolution (≤ 1.0 Å) neutron diffraction. Rather than draw a regression line as in Fig. 1a, however, we draw a limiting line below which configurations are not—within experimental error—observed. There is thus apparently an excluded region. This limiting line represents the extreme limits of H-bond bending, which can be thought of as

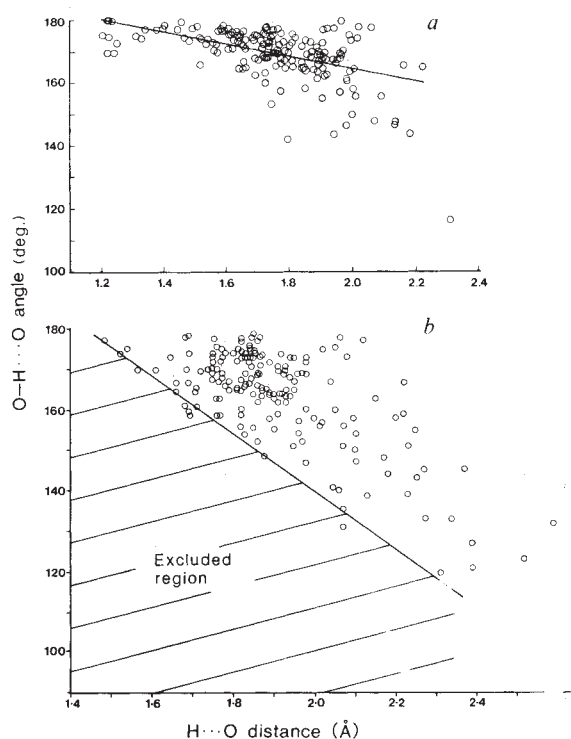
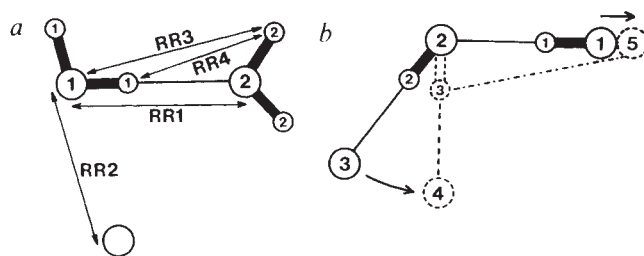


Fig. 1 Plots of O—H...O angle versus H...O distance: a, for H-bonds in small neutron structures (redrawn from ref. 3); b, for water H-bond geometries found in high-resolution neutron hydrate structures, showing the effect of repulsive restraint RR1 (see Fig. 2a) (see ref. 2 for full list of structures and source references).



characterizing the limits of the O...O repulsions of the water hydrogen-bonded neighbours. We label this 'repulsive restraint' RR1 which is illustrated, together with three other repulsive restraints, in Fig. 2a.

The second repulsive restraint, RR2, describes the significant anisotropy of O...O non-H-bonded contacts, which range from 3.1 to 3.6 Å. Figure 3a shows the minimum O...O next-nearest-neighbour contacts observed in water molecules with respect to the two orientational angles defined in Fig. 3a inset, and can be discussed in terms of three regions. In the 'lone pair' region, A, there are no contacts ≤ 3.5 Å. In the hydrogen-bonding region, B, the minimum contact distance is ~3.3 Å, whereas between the lone pairs and hydrogens (region C) a minimum separation of ~3.1 Å is observed. These data can be rationalized to a first approximation if we tentatively assign to the three separate regions around the water molecule the following van der Waals' radii: 1.7–1.8 Å in the lone-pair region, 1.6–1.7 Å between the hydrogens, and 1.5–1.6 Å between the lone pairs and the hydrogens.

The third restraint, RR3, relates to the H2...O1 remote neighbours (Fig. 2a) and appears to be the key to water molecule

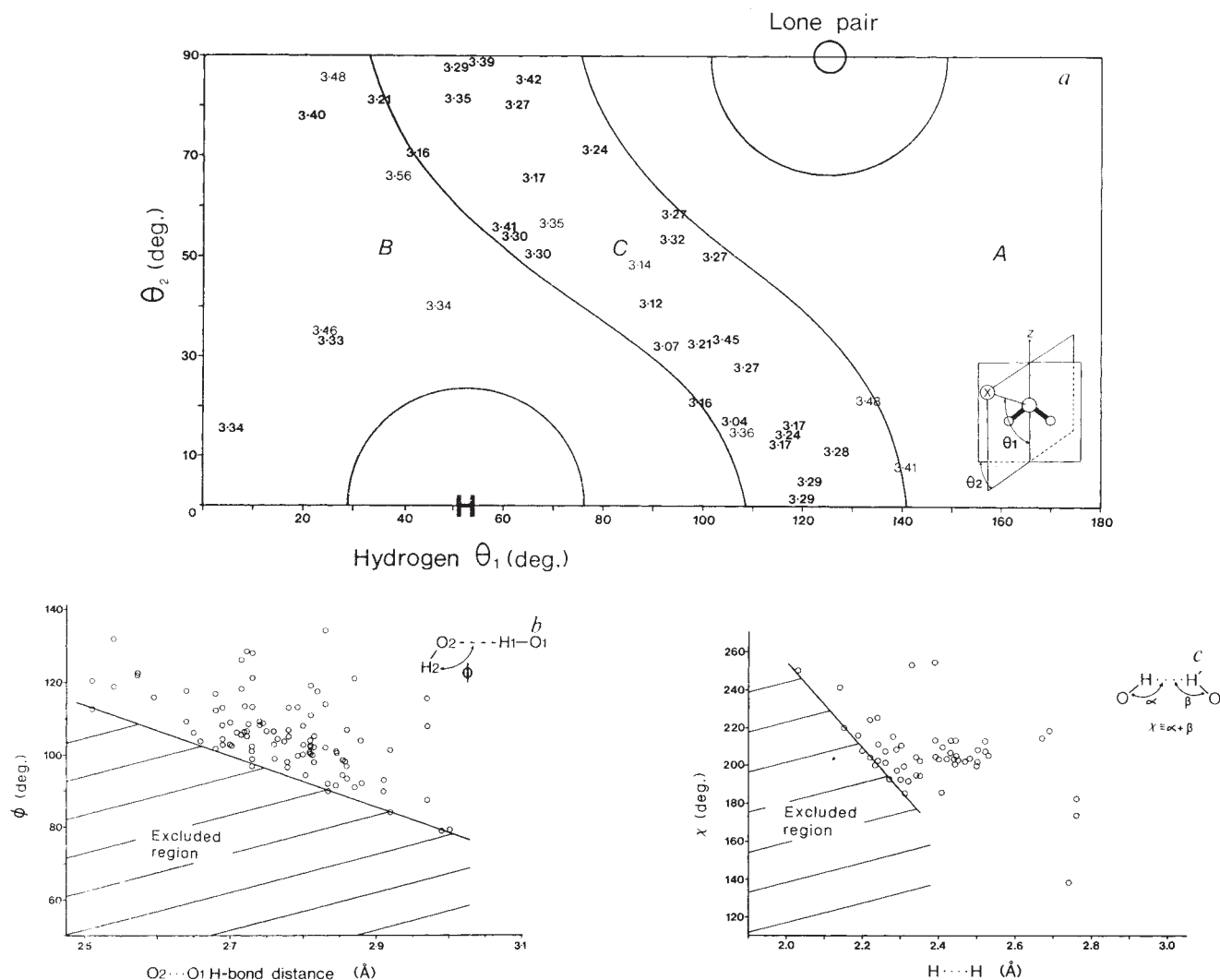


Fig. 3 a, θ_1 - θ_2 orientational angle plot for the non-bonded O...O, C contacts around four-coordinated water molecules in the neutron structures examined. Bold numbers are water-oxygen contact distances (in Å) to oxygens; lighter numbers to carbons. θ_1 and θ_2 are the angles of approach of the contact, as defined in the inset. The significance of regions A, B and C to RR2 is discussed in the text. b, Plot of H2-O2...O1 angle (ϕ) versus O2...O1 H-bond distance, showing the effect of repulsive restraint RR3, relating to H2...O1 contacts (see Fig. 2a). In this part of the analysis, the O-H covalent bonds were standardized to 0.8 Å, to enable the approximate centre of the electron density of the hydrogens to be represented. This distance, the numerical value of which is not critical, was chosen on the basis of results from high-resolution X-ray structure analyses^{6,7} and deformation studies⁸⁻¹⁰. c, Plot of the sum (χ) of the O-H...H angles subtended at the hydrogens (see inset) versus the H...H contact distance, showing the effect of repulsive restraint RR4 (see Fig. 2a).

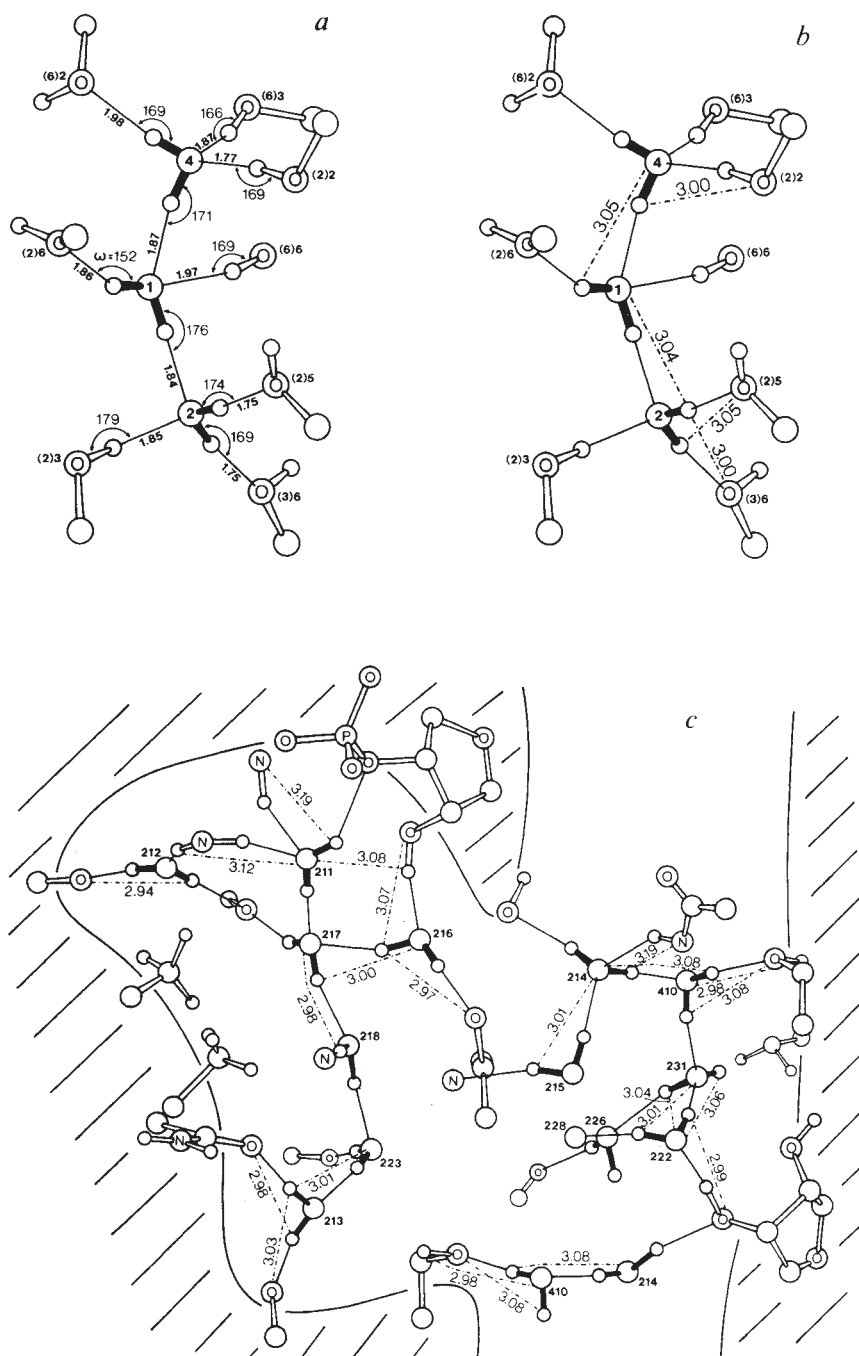


Fig. 4 *a, b*, Hydrogen-bond networks in a local region of α -cyclodextrin, showing *a*, the range of H-bond distances (in Å) and particularly H-bond angles (in degrees); and *b*, an explanation of the orientations on the basis of minimum oxygen-remote-hydrogen (H2...O1) contacts (shown in Å) of ~ 3.0 Å (effect of RR3). *c*, H2...O1 contacts (in Å) around water molecules in vitamin B₁₂ co-enzyme hydrate, shown in projection.

orientations. Figure 4*a* shows a region in α -cyclodextrin⁴. The hydrogen-bond distances and angles vary widely within the range shown in Fig. 1. Using the regression line as a target geometry, we would expect the H-bond angle ω (152°) to be greater, as a largely rotational motion of water 1 could be made without significantly affecting the near-linear O1—H...O2 angle. However, experimentally, this does not happen. Figure 4*b* shows the same region, but with the relevant remote-neighbour H2...O1 distances given, and immediately we may understand the situation: no remote-neighbour H...O distances are found which are ≤ 3.0 Å. Viewed in this way, an attempt to increase ω would begin to compromise the presumably repulsive interaction with the oxygen of water 4.

Figure 4*c* shows the H2...O1 minimum contacts between water molecules in the crystal hydrate of co-enzyme B₁₂ (ref. 5

and J.L.F., P. F. Lindley, H.F.J.S. and P. A. Timmins, in preparation), and the same pattern is apparent. These minimum contacts appear to 'fix' the mutual orientations of the water molecules at the expense of distortions of the H-bond geometries. In Fig. 3*b*, the H2—O2...O1 angle is plotted against the O2...O1 H-bond distance. As in Fig. 1, rather than draw a regression line to represent the mean values, a line representing the minimum values of the H2...O1 contacts (H2...O1 limiting curve) is drawn to give another section through what is presumably an excluded volume of the total configurational space.

The fourth constraint, RR4, relates to the H...H repulsions and Fig. 3*c* shows a plot of the water H...H repulsive contacts as a function of the angle χ (defined as the sum of the two angles subtended at the hydrogens, see Fig. 3*c* inset). The shortest water H...H contacts are ~ 2.1 Å and occur when the

O—H groups are arranged more nearly 'head-on': the majority are between 2.3 and 2.4 Å. Again, an excluded region can be specified.

Using these regularities, we can now understand more clearly some structural tendencies shared by ices and hydrates. For example, there are two ways of obtaining a more compact water structure than in ice Ih: by decreasing some of the O···O···O H-bond angles and/or by forming interpenetrating H-bond lattices (for example, at high pressures). By the former method, some of the O···O H-bond distances must increase in order to maintain the minimum H2···O1 contacts (Fig. 2b). These contacts appear to play a dominant role in determining the immediate local water structure. For example, acceptor waters, which form very short H-bonds (down to ~2.5 Å), tend towards a trigonal planar coordination, in order to maintain the H2···O1 minimum contacts of ~3.0 Å (ref. 2). Where longer H-bonds of 2.8–3.2 Å occur, significant distortions from the expected tetrahedral structure are allowed. Thus, in water structures in which the number and strength of the H-bonds are maximized (shorter H-bonds), the H2···O1 minimum contacts tend to enforce a tetrahedral-like structure (for example, in the ice polymorphs).

The exact nature of these interactions is not clear in all cases; however, we can make some proposals which may assist in developing better potential functions for simulation calculations. RR2 can be considered in terms of an angular-dependent van der Waals' radius which is, not surprisingly, larger around the lone-pairs region. The RR3 and RR4 interactions may be understood in terms of assigning a variable, angular-dependent van der Waals' radius to the polar water-hydrogens. Values of ~1.5 Å in the H—O direction and ≤1.0 Å in the O—H direction may be appropriate. This type of model is consistent with the known asphericity of the electron cloud around hydrogens attached to polar groups^{6,7}. More exact values may be obtained from more detailed analyses of the interactions between electron clouds of neighbouring water molecules (for example, by performing detailed high-level quantum mechanical calculations).

In the absence of a full physical explanation of the nature of these interactions, the very clear pattern of excluded regions, with no known significant violations, gives us confidence in using these minimum values as strict constraints in interpreting water organization in crystals (for example, during protein refinement). The 'hardness' of these excluded regions makes them far more reliable than the softer, more indefinite bond-angle and bond-length restraints which have been used until now and which, as shown in Fig. 4a, b, can lead to incorrect structures.

The repulsive restraints discussed here can also be used to improve potential functions widely used in computer simulations of aqueous systems. A reasonable model for the non-bonded intermolecular structure will have to take into account the asymmetric character of both the oxygens and the hydrogens; that this has not been done may account for the difficulty in obtaining a satisfactory agreement with experiment¹. We are pursuing the incorporation of repulsive restraints in water potentials for the simulation of liquid water and other chemically and biologically important aqueous systems.

Received 11 April; accepted 7 May 1986.

1. Finney, J. L., Quinn, J. E. & Baum, J. O. in *Water Science Reviews* Vol. 1 (ed. Franks, F.) 93–170 (Cambridge University Press, 1985).
2. Savage, H. F. J. in *Water Science Reviews* Vol. 2 (ed. Franks, F.) (Cambridge University Press, in the press).
3. Olovsson, I. & Jönsson, P. G. in *The Hydrogen Bond* Vol. 2 (eds Schuster, P., Zündel, G. & Sandorfy, C.) 393 (North-Holland, Amsterdam, 1976).
4. Klar, B., Hingerty, B. & Saenger, W. *Acta crystallogr.* B36, 1154–1165 (1980).
5. Savage, H. F. J. *Biophys. J.* (in the press).
6. Hamilton, W. C. & Ibers, J. A. *Hydrogen Bonding in Solids* (Benjamin, New York, 1968).
7. Bacon, G. E. *Neutron Diffraction* 355 (Clarendon, Oxford, 1975).
8. Savariault, J. M. & Lehmann, M. S. *J. Am. Chem. Soc.* 102, 1298–1303 (1980).
9. Dam, J., Harkema, S. & Feil, D. *Acta crystallogr.* B39, 760–768 (1983).
10. Hermansson, K. & Thomas, J. O. *Acta crystallogr.* C39, 930–936 (1983).

Selenium in western Atlantic precipitation

Gregory A. Cutter* & Thomas M. Church†

* Department of Oceanography, Old Dominion University, Norfolk, Virginia 23508, USA

† College of Marine Studies, University of Delaware, Newark, Delaware 19711, USA

The trace element composition of precipitation (rain and snow) provides a way of monitoring anthropogenic and natural emissions to the atmosphere; the subsequent deposition of such emissions (for example, as acid precipitation) can have important effects on aquatic ecosystems. In this regard, the mobilization of selenium by fossil fuel combustion¹ and selenium's toxicity² provide compelling reasons to examine this element's atmospheric transport and removal. Furthermore, recent investigations suggest that atmospheric deposition is an important selenium input to oceanic surface waters^{3,4}. Here we examine data on the concentration and oxidation state of selenium in precipitation over coastal and mid-ocean regions of the western Atlantic Ocean. The results indicate that fossil fuel combustion enriches selenium in wet depositional fluxes to western Atlantic surface waters, and that selenium's oxidation state in precipitation may be a sensitive oxidation-reduction (redox) tracer.

Selenium is removed from the atmosphere by processes such as wet deposition, through its association with sub-micrometre aerosols⁵ and the efficient scavenging of these small particles during precipitation events⁶. As a Group VI element, selenium can exist in four oxidation states (–II, selenide; 0, elemental selenium; IV, selenite; VI, selenate). Selenium is thought to be naturally emitted to the atmosphere as volatile dimethyl selenide⁷; selenium dioxide [Se(IV)] or elemental selenium may be released during fossil fuel combustion⁸. In rainwater samples from Japan⁹ and coastal California¹⁰, selenite is the major selenium species; rain and snow in urban Belgium contain variable quantities of both selenite and selenate¹¹. Further studies of selenium's chemical speciation in precipitation may help to identify the sources of atmospheric selenium, elucidate the equilibrium state and kinetics of atmospheric redox reactions, and quantify the contribution of atmospheric deposition to the oceanic selenium cycle.

Precipitation samples were obtained at Lewes, Delaware and High Point, Bermuda, from August 1983 to January 1984 using automated Aerochem Metrics precipitation samplers¹². Rain water was placed in acid-cleaned borosilicate bottles, acidified to pH 2 (with HCl), and refrigerated to minimize biological growth. Determinations of selenite, selenate, and selenide + elemental selenium were made using the procedures described in refs 3, 10 and 13; sulphate, sodium and pH were determined using the methods in refs 14 and 15. To evaluate possible artefacts arising from sample collection, two precipitation events were sampled using three different collection devices; the concentration and speciation of selenium from all three systems were within 10% of each other. In addition, artefacts from sample storage appeared to be minimal, as no selenium speciation changes were detected between rain water analysed within hours of collection and that analysed one month later. Thus, the data presented below are assumed to be representative of atmospheric processes and not procedural aberrations.

The sampling site at Lewes receives precipitation which is fairly representative of northeastern United States air masses¹⁶, and winter storm fronts often pass over Bermuda from North America after 1–2-day transit times¹⁷. Thus, the rain water concentration and speciation of selenium collected at these locations may reflect the combined effects of source strengths (marine and continental), and processes which occur during atmospheric transport and deposition. Figure 1a is a plot of total selenium concentration versus sodium at Lewes and Bermuda. The

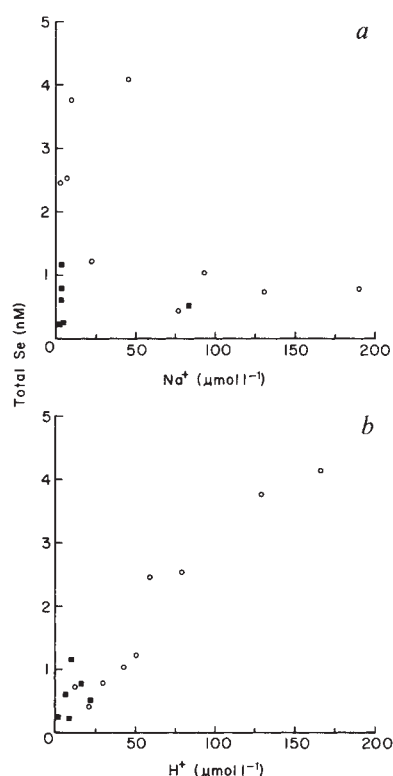


Fig. 1 Total selenium versus sodium (*a*) and protons (*b*) in precipitation samples collected at Lewes, Delaware (○) and Bermuda (■). Values for total selenium are based on triplicate determinations, with the procedural precision (expressed as relative standard deviation) not exceeding 3%. The correlation coefficient (*r*) for a linear regression fit of the total selenium/proton data is 0.9589 (*n* = 15).

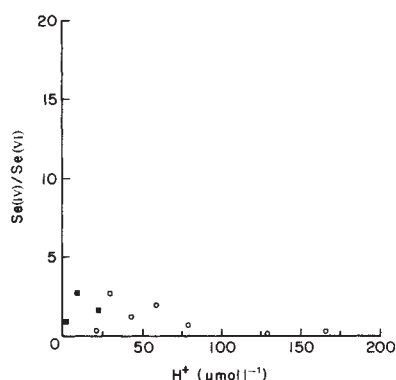


Fig. 2 The ratio of selenite to selenate [Se(IV)/Se(VI)] versus protons in Lewes, Delaware (○) and Bermuda (■) precipitation samples. Determinations of selenite and selenate were made in triplicate, and the procedural precision (expressed as relative standard deviation) did not exceed 7%.

apparent lack of correlation between selenium and sodium indicates that processes generating sea-salt aerosols do not contribute selenium to precipitation. However, there appear to be two distinct groups of precipitation events, those with low selenium-high sodium (marine origin) and those with high selenium-low sodium (non-marine). The low selenium-concentration background is in accord with the presence of uniform concentrations of aerosol selenium in the global marine atmosphere⁵. In contrast, Fig. 1*b* shows a strong correlation between total selenium and protons. A similar correlation is found between total selenium and excess (non-sea-salt) sulphate (*r* = 0.9505, *n* = 15). Previous studies of western Atlantic precipitation have attributed increased acidity and excess sulphate to

anthropogenic emissions from North America^{16,18}. The data shown in Fig. 1*b* suggest that fossil fuel combustion in North America also contributes selenium to western Atlantic precipitation.

In Fig. 2, the ratio of selenite to selenate in precipitation is plotted against proton concentration. During this study, selenide + elemental selenium was detected in only one sample (Lewes; 28% of the total selenium). Furthermore, with the exception of two samples from Lewes, the range of selenite/selenate values at both sites is small [Se(IV)/Se(VI) = 1.26 ± 0.95, *n* = 10]. The two samples with the highest selenite/selenate ratios came from storms travelling in northerly directions, placing the coal-fired Indian River Station power plant 30 km upwind of the rain sampler. Thus, a local anthropogenic input may account for the enrichment of selenite in these two events. For the remaining samples, the source-type (that is, natural or anthropogenic) cannot be identified on the basis of oxidation state in Fig. 2. If the oxidation states of atmospheric selenium inputs are correctly assigned (as above), then the rates of redox reactions during transport and deposition must be sufficiently rapid to obscure discrete oxidation state signals.

If the results in Fig. 2 are generally indicative of oxidation-reduction equilibrium, then the apparent redox intensity of precipitation, *pE* (where *pE* = -log [*e*⁻]), can be estimated using the equation: $\text{SeO}_4^{2-} + 3\text{H}^+ + 2\text{e}^- = \text{HSeO}_3 + \text{H}_2\text{O}$; *E*⁰ = 1.075 V. The *pE* of 11.50 ± 1.01 (*n* = 10) calculated in this manner is close to values for the nitrate-nitrogen dioxide and peroxide-oxygen couples reported elsewhere¹⁹. However, more data are needed on the speciation of selenium, as well as other redox couples, in precipitation, to determine more accurately the oxidation intensity of corresponding air masses.

The volume-weighted average selenium concentration (that is, the total selenium normalized to rainfall amount) at Bermuda is <20% of that at Lewes (Table 1). However, the washout

Table 1 Selenium in western Atlantic precipitation

Volume-weighted average total selenium (pmol cm ⁻³)	
Bermuda	0.38
Lewes	2.20
Average rainfall (cm yr ⁻¹)	
Bermuda	150
Lewes	90
Average total selenium in air (pmol m ⁻³)	
Bermuda ²⁰	2.4
Coastal Thode Island ⁵	11.8
Washout ratio (vol-wtd avg./air conc.)	
Bermuda	2,000
Lewes (using Rhode Island air conc.)	2,365
Wet depositional flux (pmol cm ⁻² yr ⁻¹)	
Bermuda	57
Lewes	190
Average upwelling flux, central Atlantic* (pmol cm ⁻² yr ⁻¹)	264

* Assumes total sub-surface Se = 0.66 nmol l⁻¹ (ref. 23); average upwelling rate = 400 cm yr⁻¹ (ref. 24).

ratios (volume-weighted average concentration in precipitation divided by concentration in air) at the two sites are equally large, suggesting that removal by wet deposition predominates. The wet depositional flux of selenium at Bermuda (Table 1) is similar to the average total deposition estimated from aerosol measurements (38 pmol cm⁻² yr⁻¹) by other workers²⁰ for the same site. Although no independent data are available, the calculated wet depositional flux at Lewes (Table 1) falls between the values estimated by Ross²¹ for remote continental (75 pmol cm⁻² yr⁻¹) and "urban" (840 pmol cm⁻² yr⁻¹) wet deposition. In spite of the large reduction in selenium flux between the coastal United States and Bermuda, this atmospheric input to the central Atlantic is still significant (22% of that due to vertical upwelling; see Table 1). These flux data

support the hypothesis by Measures *et al.*⁴ that atmospheric deposition elevates surface-water selenium concentrations in the western Atlantic. Furthermore, selenium is enriched in precipitation over the western Atlantic relative to sub-surface waters (average sub-surface selenite/selenate = 0.09, ref. 21). As a result, inputs of selenite and selenate through wet deposition and upwelling have different effects on parameters such as surface-water residence times for these two species³.

The selenium speciation data obtained for wet deposition suggest that the element may act as a useful tracer of oxidation-reduction equilibrium and kinetics. More thorough sample collection and analysis efforts will be required to verify the extent of this equilibrium. In addition, the sequential sampling of

individual storm events at the coast and at multiple sites across the Atlantic should yield data on the kinetics of oxidation-reduction reactions affecting selenium. The chemical similarity between selenium and sulphur increase the interest of such studies. The equilibrium between sulphur dioxide and sulphuric acid is affected by the atmosphere's redox intensity (*pE*). The rate at which this equilibrium is reached controls the long-range atmospheric transport of sulphur, the deposition of sulphuric acid, and thus affects acid precipitation²².

This is a contribution to the WATOX program, supported by a grant from the NOAA Air Resources Laboratory. We thank J. Scudlark for sample collection and L. Cutter for sample analysis.

Received 24 February; accepted 8 May 1986.

- Bertine, K. K. & Goldberg, E. D. *Science* **173**, 233-235 (1971).
- Shamberger, R. J. *The Biochemistry of Selenium*, 185 (Plenum, New York, 1983).
- Cutter, G. A. & Bruland, K. W. *Limnol. Oceanogr.* **29**, 1179-1192 (1984).
- Measures, C. I., Grant, B., Khadem, M., Lee, D. S. & Edmond, J. M. *Earth planet. Sci. Lett.* **71**, 1-12 (1984).
- Mosher, B. W. & Duce, R. A. *J. geophys. Res.* **88**, 6761-6768 (1983).
- Jickells, T. D., Knap, A. H. & Church, T. M. *J. geophys. Res.* **89**, 1423-1428 (1984).
- Jiang, S., Robberecht, H. & Adams, F. *Atmos. Envir.* **17**, 111-114 (1983).
- Andren, A. W., Klein, D. J. & Talmi, Y. *Envir. Sci. Technol.* **9**, 856-858 (1975).
- Suzuki, Y., Sugimura, Y. & Miyake, Y. *J. met. Soc. Japan* **59**, 405-409 (1981).
- Cutter, G. A. *Analyt. chim. Acta* **98**, 59-66 (1978).
- Robberecht, H. & Van Grieken, R. in *Trace Substances in Environmental Health* Vol. 14 (ed. Hemphill, D. D.) (University of Missouri, Columbia, 1980).
- Tramontano, M., Scudlark, J. S. & Church, T. M. *Environ. Sci. Technol.* (submitted).
- Cutter, G. A. *Analyt. chim. Acta* **149**, 391-394 (1983).
- Galloway, J. N., Crosby, B. J. Jr. & Likens, G. E. *Limnol. Oceanogr.* **24**, 1161-1165 (1979).
- Galloway, J. N. & Likens, G. E. *Tellus* **30**, 71-82 (1978).
- Church, T. M., Galloway, J. N., Jickells, T. D. & Knap, A. H. *J. geophys. Res.* **87**, 11013-11018 (1982).
- Miller, J. M. & Harris, J. M. *Atmos. Envir.* **19**, 409-414 (1985).
- Jickells, T. D., Knap, A. H., Church, T. M. & Galloway, J. N. *Nature* **297**, 55-57 (1982).
- Morgan, J. J. in *Atmospheric Chemistry* (ed. Goldberg, E. D.) (Springer, New York, 1982).
- Duce, R. A. *et al.* in *Marine Pollutant Transfer* (eds Windom, H. L. & Duce, R. A.) (Lexington Books, Lexington, 1976).
- Ross, H. B. *Tellus* **37**, 78-90 (1985).
- Rohde, H., Crutzen, P. & Vanderpol, A. *Tellus* **33**, 132-141 (1981).
- Measures, C. I. & Burton, J. D. *Earth planet. Sci. Lett.* **46**, 385-396 (1980).
- Broecker, W. S. & Peng, T. H. *Tracers in the Sea* (Eldigio, New York, 1982).

Correlations between stream sulphate and regional SO₂ emissions

Richard A. Smith & Richard B. Alexander

United States Geological Survey, Reston, Virginia 22092, USA

The relationship between atmospheric SO₂ emissions and stream and lake acidification has been difficult to quantify, largely because of the limitations of sulphur deposition measurements. Precipitation sulphate (SO₄) records are mostly <5 yr in length¹ and do not account for dry sulphur deposition². Moreover, a variable fraction of wet- and dry-deposited sulphur is retained in soils and vegetation and does not contribute to the acidity of aquatic systems^{3,4}. We have compared annual SO₂ emissions for the eastern United States from 1967 to 1980 with stream SO₄ measurements from fifteen predominantly undeveloped watersheds (Figs 1, 2). We find that the two forms of sulphur are strongly correlated on a regional basis and that streams in the southeastern United States (SE) receive a smaller fraction (on average, 16%, compared with 24%) of regional sulphur emissions than do streams in the north-eastern United States (NE). In addition to providing direct empirical evidence of a relationship between sulphur emissions and aquatic chemistry, these results suggest that there are significant regional differences in the fraction of deposited sulphur retained in basin soils and vegetation.

The stream SO₄ data (Table 1) are from monitoring stations in the USGS Hydrologic Benchmark Network⁵. Five stations were used to estimate average annual SO₄ yield in the NE, and 10 stations were used in the SE (Fig. 1). Stream SO₄ at these stations is believed to be largely of atmospheric origin^{6,7}. The SO₄ yield for each year from 1967 to 1980 was estimated as the product of the annual discharge-weighted mean SO₄ concentration and mean basin runoff for the 14-yr period. The regional SO₄ yield was estimated as the average of the individual basin yields.

Regional SO₂ emission densities were based on state-level data from two recent sources^{8,9} which used alternative methods of estimating emissions. Regressions against stream sulphate were conducted separately for the two sets of emission data in order to determine the sensitivity of results to the choice of methods. Husar's⁸ 'production-driven' estimates are based on

records of the sulphur content and distribution of produced fuels, whereas those from Gschwandtner *et al.*⁹ are 'consumption-driven' and are based on fuel consumption records. Emission estimates from Hidy *et al.*¹⁰ for southeastern Ontario were added to both data sets in calculating NE emission densities and represent ~17% of the regional total. Trends in state-level SO₂ emissions^{8,9} and stream SO₄ concentrations (Table 1) were nearly homogeneous in direction (decreasing in the NE and increasing in the SE) within the defined regions during the study period. Varying the size of the defined regions through the addition or deletion of states near the regional boundaries has relatively little effect on regional emission densities, which, like stream yields, are expressed as areal averages (g m⁻² yr⁻¹).

Regressions of stream SO₄ yield against SO₂ emission density give slope values ranging from 0.14 to 0.25 for various combinations of region and source of emission data (Fig. 2). The results are similar for the two sets of emission data and suggest that,

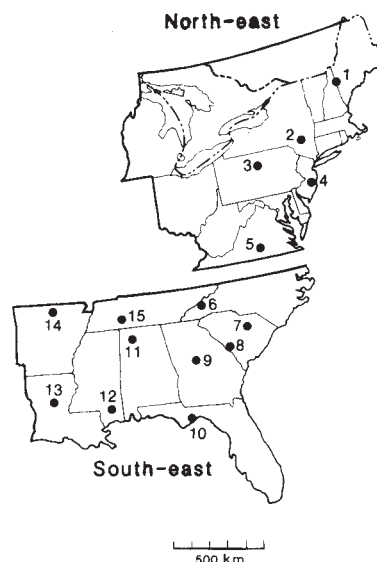


Fig. 1 Locations of stream sampling stations in the northeastern and southeastern United States. The numbered stations are identified in Table 1.

Table 1 Stream SO₄ data and wet SO₄ deposition estimates for Benchmark stream monitoring sites in the eastern United States

Station	Basin area (km ²)	Mean discharge (m ³ s ⁻¹)	Avg. stream SO ₄ concentration (μeq l ⁻¹)	SO ₄ trend ⁷	Mean SO ₄ * yield, Y (g m ⁻² yr ⁻¹)	Wet SO ₄ * deposition ¹³ , D (g m ⁻² yr ⁻¹)	Y/D
North-east region (NE)							
1. Wild River at Gilead, Maine	180	4.56	102.7	—	1.31	0.86	1.51
2. Esopus Creek at Shandaken, New York	154	3.79	162.1	—	1.95	1.16	1.68
3. McDonalds Branch in Lebanon State Forest, New Jersey	6	0.06	154.6	—	0.75	1.08	0.69
4. Young Woman's Creek near Renovo, Pennsylvania	120	2.04	164.0	—	1.44	1.16	1.24
5. Holiday Creek near Andersonville, Virginia	22	0.23	60.6	—	0.34	1.00	0.34
South-east region (SE)							
6. Cataloochee Creek near Cataloochee, North Carolina	127	3.17	25.6	+	0.35	1.09	0.32
7. Scape Ore Swamp near Bishopville, South Carolina	249	2.77	63.0	+	0.36	0.89	0.40
8. Upper Three Runs near New Ellenton, South Carolina	225	3.11	21.7	+	0.14	0.88	0.16
9. Falling Creek near Juliette, Georgia	187	1.50	101.1	0	0.44	0.81	0.55
10. Sopchoppy River near Sopchoppy, Florida	264	4.44	94.4	+	0.90	0.91	0.99
11. Sipsey Fork near Grayson, Alabama	233	4.41	85.7	+	0.81	0.95	0.85
12. Cypress Creek near Janice, Mississippi	135	2.72	44.4	+	0.49	0.73	0.67
13. Big Creek at Pollock, Louisiana	132	1.47	47.4	+	0.29	0.68	0.42
14. North Sylamore Creek near Fifty Six, Arkansas	150	1.19	119.1	+	0.47	0.89	0.53
15. Buffalo River near Flat Woods, Tennessee	1,157	21.34	84.2	—	0.85	1.16	0.74

* Expressed as sulphur.

an average, streams transport ~20% of sulphur emissions in the eastern United States. However, an analysis of covariance¹¹ indicates that, for both sets of emission data, a regression model with a single intercept but regionally separate slope estimates is preferable to other models. Slope estimates for both data sets show significantly higher values for the NE region than for the SE (an average of 24%, compared with 16%). All four regression lines pass close to the origin, a result which supports the assumption that stream SO₄ at these sites is largely of atmospheric origin and comes primarily from sources within the region.

Johnson *et al.*¹² have suggested that streams in the southeastern United States might be expected to transport a smaller fraction of deposited sulphur than northeastern streams, on the basis of experiments showing greater sulphur retention in soils typical of the SE. Southeastern soils are predominantly ultisols, which contain subsurface clay horizons with high SO₄-adsorption capacities¹². The hypothesis of greater SO₄ retention in the SE is strongly supported by the regional differences in regression slopes observed in this study, together with SO₄

deposition measurements¹³ for the regions. Comparison of stream SO₄ yields with wet SO₄ deposition estimates (Table 1) gives significantly larger ratios of yield to deposition (Y/D) in the NE. Ratios for the SE are all <1.0, whereas those for the NE are mostly >1.0. The sulphur-adsorbing properties of basin soils are thought to be an important factor controlling the acidification of surface waters because mobility of sulphate in soils is required for the transport of hydrogen ions (as well as base cations) through basin soils to streams^{3,14}.

The regression results and yield-to-deposition ratios presented above can be used to calculate approximate sulphur budgets for the NE and SE (Table 2). Regression slopes for the two regions indicate that 16% of the sulphur emitted in the SE (24% in the NE) is transported by streams, and that the remaining 84% (76% in the NE) is either retained in basin soils and vegetation or is deposited outside the region. If dry deposition is assumed to be equal to wet deposition¹⁵, then median basin retention of SO₄ is 35% of emissions in the NE and 59% of emissions in the SE. Basin retention is considerably smaller,

Table 2 Estimated sulphur budgets for the northeastern and southeastern United States (1980)

Budget component	Regional estimates			
	NE	SE	NE	SE
	Quantity (g m ⁻² yr ⁻¹)	Quantity (g m ⁻² yr ⁻¹)	Per cent of emissions	Per cent of emissions
SO ₂ emissions	3.7	2.4	100	100
Atmospheric transport out of the region	1.5	0.6	41	25
Total deposition	2.2	1.8	59	75
Basin retention	1.3	1.4	35	59
Stream transport	0.9	0.4	24	16

Total deposition is calculated as twice the measured¹³ wet deposition, atmospheric transport of sulphur from the region as the difference between SO₂ emissions and total deposition, and basin retention as the difference between stream SO₄ transport and total deposition.

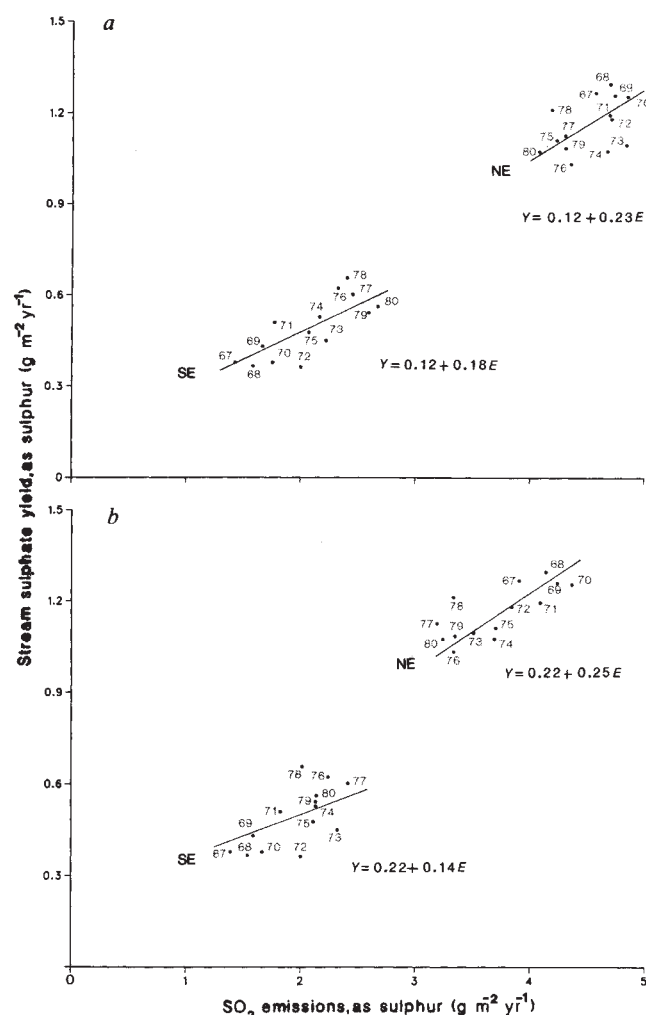


Fig. 2 Average stream yield (Y) of SO_4 at USGS Benchmark stations in the eastern United States, plotted against regional SO_2 emission density (E) for the years 1967–80 (shaded numbers). Plots are presented for two recent sets of emission estimates^{8,9}, based on alternative methods of estimating the sulphur content of combusted fuels (a, production-driven emissions⁸; b, consumption-driven emissions⁹). Both emissions and stream yield of sulphur decreased in the NE but increased in the SE during the 1967–80 period. Analyses of covariance for both sets of emission estimates result in a two-slope, one-intercept model with smaller slopes for the SE than for the NE. All slopes are significant at $\alpha = 0.01$. Intercepts are not significantly different from zero at $\alpha = 0.05$.

however, for certain of the NE basins with non-ultisolic soils. Average retention for Maine, New York and Pennsylvania, for example, is only 14%. Assuming that dry and wet deposition are equal, atmospheric transport from the region is 41% in the NE and 25% in the SE. These estimates of extra-regional transport bracket a previously reported estimate¹⁶ (30%) for sulphur transport from eastern North America, thus supporting the earlier estimate.

There has been continuing uncertainty concerning the quantitative relationship between atmospheric SO_2 emissions and sulphur deposition in North America and, to a lesser extent, in Europe^{15,17}. This uncertainty stems from the shortage of long-term deposition records on which to base correlative studies¹⁸, as well as from problems with the measurement of dry deposition. The relationship between SO_2 emissions and aquatic transport of SO_4 has remained even less clear because of a shortage of empirical data on the retention of wet- and dry-deposited

sulphur in a variety of soil types⁶. The results described here indicate that an approximately linear (and nearly proportional) relationship exists between SO_2 emissions and stream transport of atmospheric sulphur in the eastern United States. These results appear to demonstrate the response of stream SO_4 to both increases and decreases in emissions over a wide range of emission densities. There is evidence that the slope of the relation between SO_2 emissions and stream SO_4 is somewhat smaller in the SE due to greater basin retention of SO_4 . Greater SO_4 retention may initially serve to reduce the acidifying effects of SO_4 deposition in the SE, but gradual accumulation of SO_4 in soils may lead to increased stream transport of SO_4 in the future. Additional variability in the effects of acid deposition can be expected to result from differences in the ability of basins to neutralize acids by supplying base cations^{6,19}. Budgets for the environmental fate of atmospheric SO_2 emissions, including basin retention of SO_4 , could be further refined by reliable measurements of dry deposition.

We thank J. Turk, G. Oehlert and A. Johnson for helpful discussions, and O. Bricker, H. Lins and J. Pickering for reviewing the manuscript.

Received 24 February; accepted 9 May 1986.

1. Stensland, G. J., Whelpdale, D. M. & Oehlert, G. in *Acid Deposition: Long-Term Trends*, 128–199 (National Academy Press, Washington, 1986).
2. Hicks, B. B. (ed.) *Deposition Both Wet and Dry* Vol. 4 (Butterworth, Boston, 1984).
3. Reuss, J. O. & Johnson, D. W. *J. environ. Qual.* **14**, 26–31 (1985).
4. National Research Council *Acid Deposition: Processes of Lake Acidification* (National Academy Press, Washington, 1984).
5. Cobb, E. D. & Biesecker, J. E. *U.S. geol. Surv. Circ.* No. 460-D (1971).
6. Kramer, J. R. *et al.* in *Acid Deposition: Long-Term Trends*, 231–299 (Washington Academy Press, Washington, 1986).
7. Smith, R. A. & Alexander, R. B. *U.S. geol. Surv. Circ.* No. 910 (1983).
8. Husar, R. B. in *Acid Deposition: Long-Term Trends*, 48–92 (National Academy Press, Washington, 1986).
9. Gschwandtner, G., Gschwandtner, K. C. & Eldridge, K. *Historical Emissions of Sulfur and Nitrogen Oxides in the U.S. from 1900–1980* (EPA 600/7-85-009a) (Environmental Protection Agency, Washington, 1985).
10. Hidy, G. M., Henry, R. C., Hansen, D. A., Ganesan, K. & Collins, J. J. *Air Pollut. Control. Ass.* **34**, 333–345 (1984).
11. Kleinbaum, D. G. & Kupper, L. L. *Applied Regression Analysis and Other Multivariable Methods* (Wadsworth, Belmont, 1978).
12. Johnson, D. W., Hornbeck, J. W., Kelly, J. M., Swank, W. T. & Todd, D. E. in *Atmospheric Sulfur Deposition: Environmental Impact and Health Effects* (eds Shriner, D. S., Richmond, C. R. & Lindberg, S. E.) 507–520 (Ann Arbor Science, 1980).
13. Gibson, J. H. & Baker, C. V. *1980–81 Wet-Deposition Measurements* (National Atmospheric Deposition Program, Colorado State University, Ft. Collins, 1982).
14. Cosby, B. J., Hornberger, G. M., Galloway, J. N. & Wright, R. F. *Wat. Resour. Res.* **21**, 51–63 (1985).
15. National Research Council *Acid Deposition: Atmospheric Processes in Eastern North America* (National Academy Press, Washington, 1983).
16. Galloway, J. N. & Whelpdale, D. M. *Atmos. Envir.* **14**, 409–417 (1980).
17. Eliassen, A. & Saltbones, J. *Atmos. Envir.* **17**, 1457–1473 (1983).
18. Oppenheimer, M., Epstein, C. B. & Yuhnke, R. E. *Science* **229**, 859–862 (1985).
19. Driscoll, C. T. & Likens, G. E. *Tellus* **34**, 283–292 (1982).

Dating ice-wedge growth in subarctic peatlands following deforestation

Serge Payette, Line Gauthier & Ivan Grenier

Centre d'études mordiques, Université Laval, Québec, Canada G1K 7P4

Ice wedges are one of the most prominent periglacial features in the zone of continuous permafrost, and constitute significant climatic and palaeoclimatic markers^{1–4}. Active ice wedges are found in areas with a mean annual air temperature below -6°C (ref. 5), although recent data⁶ suggest that they can form at -3.5°C . Most of the reports^{7–9} dealing with ice wedges have referred to Arctic and Antarctic conditions, whereas subarctic ice wedges from glaciated areas are poorly known. Ice-wedge inception and spatial development as orthogonal and polygonal nets were inferred generally from theoretical analysis^{10,11} and sparse detailed field work^{7,8}. Because ice wedges occur also in subarctic permafrost

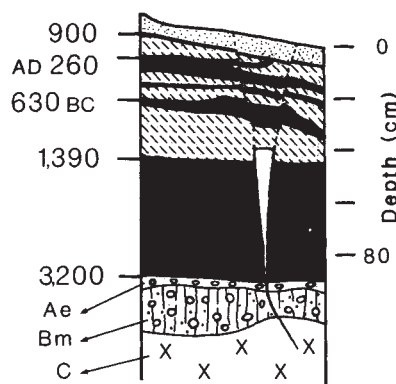


Fig. 1 Stratigraphy of a peat section from outside the studied quadrangle, showing the relationship between radiocarbon-dated wood layers and ice-wedge growth. The formation of this 8-cm-wide ice began sometime after AD 900 (extrapolated from radiocarbon-dated growth curve) when spruce (black) and *Sphagnum* (hatched) cushions, always intermingled with spruce branches and leaves, were replaced by a snow-free lichen-*Dicranum* cover (dotted). The layer from 3,200 to 1,390 BC is made of intermingled spruce and *Sphagnum*. From 3,200 BC to AD 900, this site was snow-protected because of continuous spruce-*Sphagnum* growth. Ae, Bm and C at the bottom of the peat profile refer to pedogenetic horizons.

peatlands¹²⁻¹⁵, the combined use of macrofossil peat stratigraphy and radiocarbon dating may yield useful information concerning their development. Here we present the first account of such a framework, derived from the analysis of marginal ice wedges developed in the eastern Canadian Subarctic. We show that the inception and spatial development of epigenetic ice wedges in some subarctic peatlands occurs after the natural removal of coniferous vegetation, which induces snow-free conditions, permafrost aggradation and, ultimately, deep frost cracking. We have dated ice wedges indirectly, using combined curves of rates of spruce removal and ice-wedge growth.

In North America, subarctic ice wedges are found mostly in peat plateaus and palsas bogs¹²⁻¹⁴. Their presence is generally inferred from the polygonal network which dissects the extensive, windswept lichen-heath vegetation occurring above the fossil peat. A closer look at the upper peat stratigraphy where epigenetic ice wedges develop reveals in most cases a charcoal horizon or a non-charred wood layer indicative of the former presence of coniferous vegetation. As long as the coniferous vegetation was able to maintain itself on the peatland, buffered thermal conditions prevailed in the peat substrate because of snow insulation. Removal of the coniferous cover induces snow-drift conditions and creates snow-free sites that change the overall thermal regime, favouring deeper cold penetration into the peat substrate. This close relationship between coniferous vegetation, snow conditions and thermal characteristics of substrates forms the basis of our hypothesis, which states that ice wedges found in many subarctic peatlands in eastern Canada, where continental deglaciation occurred during early to mid-Holocene times^{16,17}, were formed sometime after site deforestation.

A small insular *Sphagnum* bog in the Clearwater Lake area, subarctic Québec (56°30' N, 75°30' W; mean annual air temperature $\approx -5^\circ\text{C}$), has been used to test the hypothesis. The peat stratigraphy of the bog (Fig. 1) shows a cyclic alternation of *Sphagnum* and spruce (*Picea mariana* (Mill.) BSP.) which persisted for several thousand years¹⁸; cessation of cyclic replacement occurred during the past 2,000 years, particularly in the most exposed sites where a lichen-*Dicranum*-heath vegetation now covers the fossil peat. Ice wedges develop in these sites and form an irregular polygonal pattern at the surface. The

measurement of 26 polygonal cells separating individual ice-wedge lineaments has indicated an average dimension of $7 \times 9\text{ m}$, although many polygons were incompletely formed. We have studied in detail ice wedges developed under a lichen-*Dicranum*-heath cover in a representative section of the selected peatland. Within a $1,444\text{-m}^2$ quadrangle ($38 \times 38\text{ m}$), the ice wedges were mapped and measured (Fig. 2). Correlation between ice-wedge vertical length and width was established by direct observation in manually excavated peat sections. Ice wedges were distinguished from frost cracks, also filled with ice but confined to the active layer, by a detailed study of their melting pattern during summer 1982. The surface topography of the quadrangle was determined with an electronic level (GDD Instrumentation Inc.), and the thickness of the peat deposit (ranging from 85 to 260 cm) was evaluated by coring at several sites. The coniferous cover of low krummholz (stunted spruces) has been mapped, and the uppermost fossil-wood layer laying immediately under the lichen-*Dicranum*-heath cover was radiocarbon dated at nine different sites within the quadrangle (Fig. 2). The ^{14}C dates were calibrated to calendar dendroyears¹⁹, to determine regression rates of coniferous vegetation and ice-wedge age (Fig. 3).

Frost cracks were particularly extensive in the study area in March–April 1982 and 1983. In 1982, the cracks were $36 \pm 6\text{ cm}$ long, ($n = 133$), and the depths of the active layer in July, August and September were, respectively, $29 \pm 2\text{ cm}$ ($n = 29$), $40 \pm 5\text{ cm}$ ($n = 113$) and $42 \pm 4\text{ cm}$ ($n = 42$). In 1983, the frost cracks were slightly longer, $39 \pm 6\text{ cm}$ ($n = 52$). Their spatial pattern was irregular, sometimes linear or arcuate, but rarely polygonal (Fig. 2).

Ice wedges were found at $49 \pm 6\text{ cm}$ ($n = 153$) from the peatland surface; that is, at the present maximum depth of the active layer. The wedges were made of white foliated ice including vertically oriented air bubbles and occasionally snow (Fig. 4a). The width of the ice wedges ranged from 0.1 to 48 cm, and was

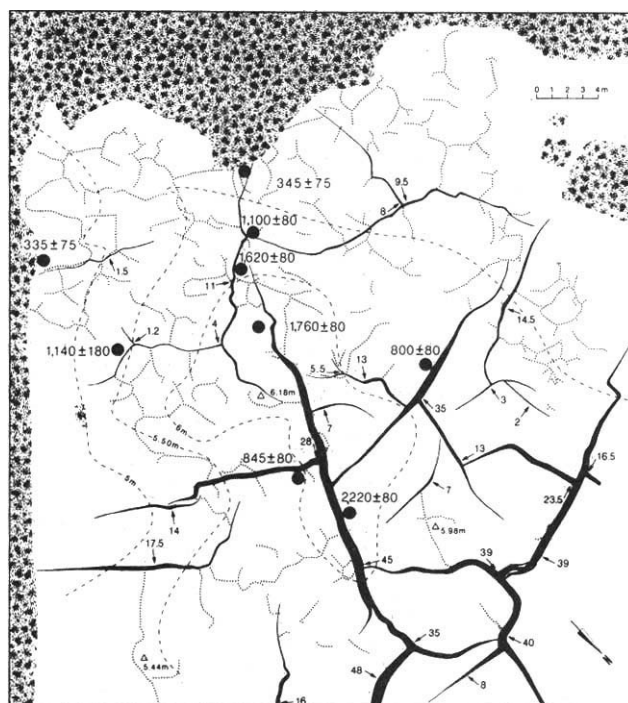


Fig. 2 Distribution of ice wedges (black) and frost cracks (dotted lines) on lichen ground cover surrounded by spruce krummholz (grey pattern). Numbered arrows denote maximum ice-wedge widths (cm). Dashed lines, contours of the peatland surface in m above lake level; open triangles, bench marks; large dots, ^{14}C dates.

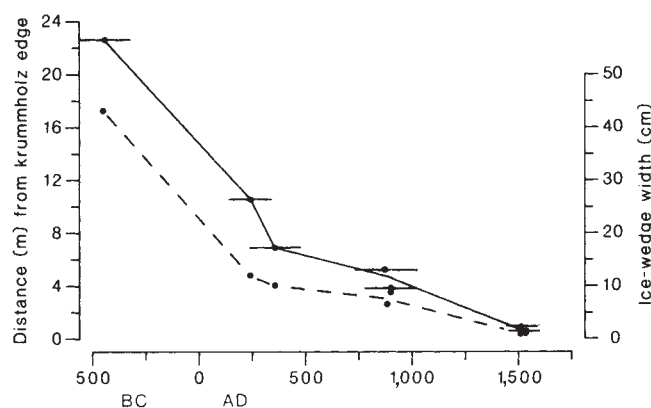


Fig. 3 Regression of coniferous vegetation (solid line), based on position of radiocarbon-dated fossil wood layers (calibrated to calendar years) relative to present krummholz edge. Horizontal bars indicate the 2σ (95% confidence) intervals for the calibrated ^{14}C dates. The curve of ice-wedge width versus time (broken line) is similar to the conifer regression curve, which gives a reliable method for ice-wedge dating in subarctic peatland ecosystems.

highly correlated ($r=0.95$) with vertical length (maximum ice-wedge length 190 cm). The ice wedges were slightly arcuate, especially at the bottom tip. Loose ice crystals (depth hoar?) were predominant at the bottom tip of large ice wedges. The basic spatial pattern consisted of individual lineaments of variable length and width, orientated in all directions. Most ice-wedge lineaments are independent features that converge occasionally in a semi-orthogonal pattern. Their horizontal growth always stops at the intersection of a nearby ice wedge, so that the longest ice wedges are those that do not merge into one another (Fig. 2). An extensive oblique layer of foliated ice formed within the mineral substrate was also observed at the intersection of two ice wedges, at a site where the peat deposit is shallower (Fig. 4b). This ice shows a structure similar to that of the wedges, an indication that its presence is associated with ice-wedge formation. Some very short but relatively wide ice

wedges have also developed, which supports the idea that horizontal growth stops when two ice wedges meet. Only two closed ice-wedge polygons were observed; they were characterized by relatively large lineaments compared to those found elsewhere.

Before ~500 BC, the peatland was entirely covered by coniferous vegetation (based on radiocarbon dating of the uppermost wood layers). Along the NE-SW direction, where two independent ice-wedge lineaments have developed in similar environments (Fig. 2), spruce regression began earlier at the higher site, and progressed radially at different rates towards the modern krummholz edges (Fig. 3). A highly significant correlation ($r=0.98$) was found between ice-wedge width and distance in metres from the modern krummholz edge; thus, the curve of maximum ice-wedge width measured near the radiocarbon-dated sites shows a pattern similar to the spruce regression curve (Fig. 3). This suggests a direct relationship between deforestation and ice-wedge inception; the age of an ice wedge is therefore very close to the age of deforestation. However, two radiocarbon dates were excluded from the analysis. The ^{14}C date of 845 ± 80 yr BP obtained from a sample close to the oldest deforested site (Fig. 2) may appear to contradict the radial pattern of spruce regression. But this date came from a charcoal sample associated with a palaeo-Indian camp fire located, probably for convenience, on the open lichen ground. All the fragments of three small boulders thought to have surrounded this fire were also recovered from the peat deposit, just above a 25-cm-wide ice wedge. This suggests that the fire was ignited from wood of nearby provenance, probably at the edge of the 850-yr BP spruce cover. Additionally, it indicates that ice-wedge activity has caused boulder burial along the thermal cracks. The ^{14}C date of 800 ± 80 yr BP appears to be anomalous according to position along the ice-wedge lineament; this may be related to the uncertain position of the sample, associated with problems of surface disturbance during peat excavation. All of the radiocarbon dates (except 845 ± 80 yr BP) confirm the radial pattern in spruce regression: from the highest site, at an altitude > 6.0 m above the Clearwater Lake level, to the modern krummholz edges, generally at an altitude < 5.0 m, spruce regression has progressed radially but at different rates according to site exposure (Fig.

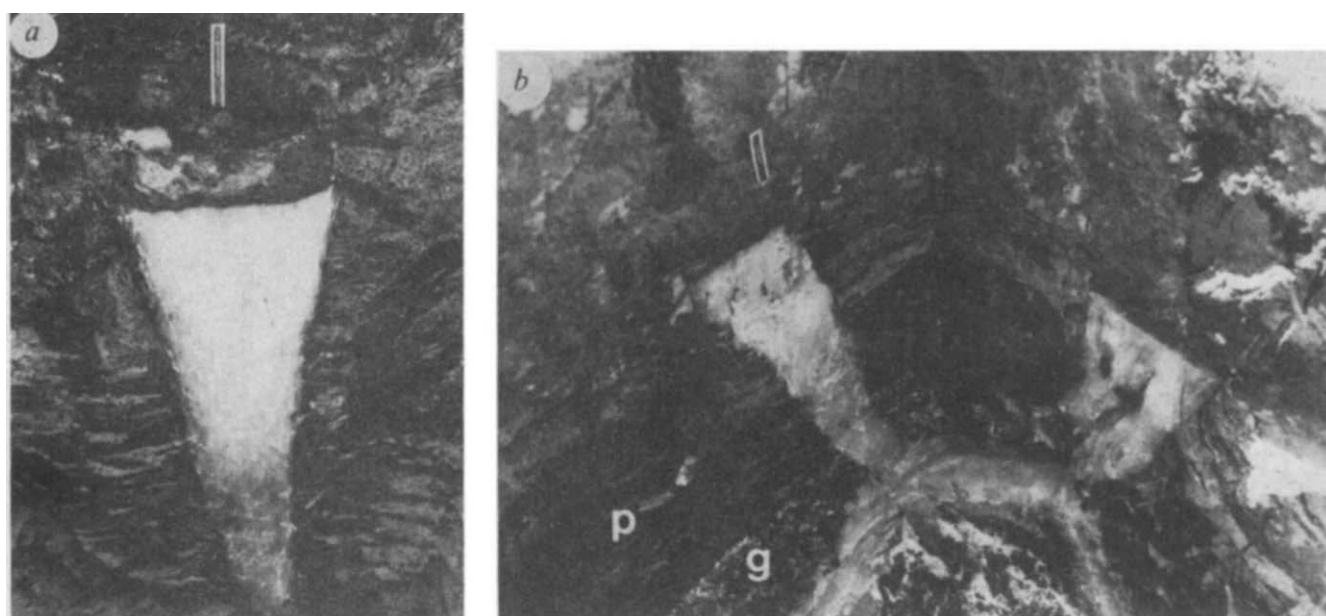


Fig. 4 *a*, Section of a 48-cm-wide ice wedge (ruler, 15 cm); ice-wedge vertical length, 190 cm. In March 1983, the left part of this ice wedge was opened by a thermal crack ~0.10–0.15 mm wide and 220 cm long. *b*, Intersection of two ice wedges, showing an extensive oblique layer of foliated ice formed within the gravelly substrate (g). The peat layer (p) is ~1 m thick.

3). The radiocarbon dates of the uppermost wood layers correspond therefore to maximum ages of ice wedges, assuming that they began to develop immediately after the removal of coniferous vegetation. Note that ice wedges do not open each winter, particularly in such marginal permafrost areas, given that in Arctic sites only 40% of the ice-wedge population is active each year⁷. In April 1983, for example, only a few ice wedges were active, although conditions for thermal contraction were particularly favourable. One 48-cm-wide ice wedge was opened by a crack ~0.10–0.15 mm wide and at least 220 cm deep (Fig. 4a); the crack was located not at the centre, but rather along one side of the ice wedge, as has frequently been observed in the study area.

The inception and spatial development of ice wedges in the subarctic requires particular microclimatic conditions that simulate the Arctic environment, where ice wedges occur extensively in mineral lowlands. These conditions are usually found in ombrotrophic peatlands, with *Sphagnum* peat as an insulating blanket. In general, ombrotrophication appears to be the necessary step for permafrost aggradation in subarctic peatlands^{12,14}, occurring some time before the microclimatic conditions for ice-wedge formation are met. In northern Canada, permafrost invasion in peatland is associated with a change in drainage conditions which is emphasized in peat stratigraphy by a shift towards ombrophilous (*Spagnum*, *Picea*, *Ericaceae*) species^{12,14}. Although regional conditions may vary²⁰, the development of epigenic ice wedges coincides generally with treeless, somewhat dried, raised peatlands characterized by a very low peat growth rate; spruce removal and expansion of windswept lichen cover seem to be the ultimate prerequisites for such a development, at least in the snowier eastern Canadian Subarctic. Subsequent ice-wedge development at these sites appears to be a slow and irregular process. Mean annual ice-wedge growth rates measured in this study vary from a maximum of 0.18 mm yr⁻¹ to a minimum of 0.04 mm yr⁻¹ (Fig. 3), a very low figure compared to modern Arctic and Antarctic ice-wedge growth rates^{7,9}. The average dimensions of these subarctic polygons are very small compared to Arctic polygons, which range from 15 to 40 m (ref. 3). Because these ice wedges have formed in a climatically marginal environment, they developed slowly, allowing us to follow the changing spatial arrangement with time. Their horizontal growth towards polygon formation did not show any defined order in time, as proposed elsewhere¹¹; in fact, the spatial growth of individual ice wedges, based on their distribution today (Fig. 2), suggests that the different segments of an ice-wedge polygon may vary in age. Whether this time-sequence in polygon formation is the rule remains to be substantiated. Our data indicate that subarctic ice-wedge expansion is a palaeoecological event closely tied to local vegetation development. Thus, the onset of ice-wedge development in the subarctic may have been directly controlled by Holocene deforestation, a long-term, time-transgressive ecological process operating at the microscale. It implies that subarctic ice wedges could have responded directly to climate forcing only when the habitat development had reached a critical threshold emphasized by deforestation. That is why the palaeoclimatic significance of subarctic ice wedges should be interpreted cautiously.

We thank Professor J. Ross Mackay (University of British Columbia) for discussion and for very useful suggestions on an earlier draft. This study has been financially supported by the Ministère de l'Éducation of Québec (FCAR Program) and the Natural Sciences and Engineering Research Council of Canada (NSERC).

5. Péwé, T. L. *Proc. 1st int. Permafrost Conf.*, 76–81 (National Academy of Science–National Research Council of Canada, Washington DC, 1966).
6. Hamilton, T. D., Ager, T. A. & Robinson, S. W. *Arct. alp. Res.* 15, 157–168 (1983).
7. Mackay, J. R. *Can. J. Earth Sci.* 11, 1366–1383 (1974).
8. Mackay, J. R. *Can. J. Earth Sci.* 12, 1668–1674 (1975).
9. Black, R. F. *Proc. 2nd int. Permafrost Conf.*, 193–203 (National Academy of Science–National Research Council of Canada, Washington DC, 1973).
10. Lachenbruch, A. *Proc. 1st int. Permafrost Conf.*, 63–71 (National Academy of Science–National Research Council of Canada, Washington DC, 1966).
11. Dostovalov, B. N. & Popov, A. I. *Proc. 1st int. Permafrost Conf.*, 102–105 (National Academy of Science–National Research Council of Canada, Washington DC, 1966).
12. Zoltai, S. C. & Tarnocai, C. *Can. J. Earth Sci.* 12, 28–43 (1975).
13. Dredge, L. A. *Geol. Surv. Pap. Can.* 79–1C, 27–30 (1979).
14. Couillard, L. & Payette, S. *Can. J. Bot.* 63, 1104–1121 (1985).
15. Dionne, J. C. *Géogr. phys. Quat.* 37, 127–146 (1983).
16. Allard, M. & Seguin, M. K. *Géogr. phys. Quat.* 39, 13–24 (1985).
17. Richard, P. J. H., Larouche, A. & Bouchard, M. A. *Géogr. phys. Quat.* 36, 63–90 (1982).
18. Payette, S. *Ecology* (submitted).
19. Klein, J., Lerman, J. C., Damon, P. E. & Ralph, E. K. *Radiocarbon* 24, 103–150 (1982).
20. Ovenden, L. *Boreas* 11, 209–224 (1982).

Hydrocarbon shows and petroleum source rocks in sediments as old as 1.7×10^9 years

M. J. Jackson, T. G. Powell, R. E. Summons & I. P. Sweet

Division of Continental Geology, Bureau of Mineral Resources, Geology and Geophysics, GPO Box 378, Canberra, ACT 2601, Australia

We report the discovery of indigenous 'live' oil in ~1.4-Gyr-old rocks in the McArthur Basin of northern Australia. Previously reported occurrences of indigenous Precambrian oil are <1 Gyr old^{1–5}. Potential petroleum source rocks in the McArthur Basin range in age from 1.4 to 1.7 Gyr and were deposited in marine and lacustrine environments. In parts of the basin they have been buried sufficiently deeply to have generated hydrocarbons. They span the period corresponding to the appearance of eukaryotic organisms, and because of their low degree of thermal alteration, they provide a valuable resource for the study of primitive biota through their hydrocarbon biomarkers. The hydrocarbon composition of the oil is consistent with a derivation from organic matter of prokaryotic origin. These results show that exploration of previously ignored mid-Proterozoic sediments may lead to the discovery of new reserves of oil.

Despite the occurrence of commercial Proterozoic oil and gas fields in the Soviet Union, Oman and China^{1–5}, rocks of this age are not generally considered to be likely sources of petroleum⁶. The McArthur Basin (Fig. 1a) contains an unmetamorphosed, structurally simple sedimentary sequence. A sequence of stromatolitic and evaporitic carbonates with interbedded shales (McArthur and Nathan groups) is overlain unconformably by quartz arenites and interbedded shales (Roper Group)⁷. The sequence ranges in age from 1,690⁺²⁹₋₂₅ Myr for the Barney Creek Formation⁸ of the McArthur Group to 1,429 ± 31 Myr for the McMinn Formation⁹ of the Roper Group (Fig. 1b). The presence of extensive unmetamorphosed carbonate shales, widespread microbial remains^{10–15} and reports of bitumen and gas in the McArthur Group¹⁶ prompted an investigation of the hydrocarbon potential of the basin.

Organic-rich sediments, at various stages of hydrocarbon generation (maturation) have been found at five different stratigraphic levels (Fig. 1b, Table 1). McArthur Group sediments were deposited in lacustrine and perhaps very shallow marginal marine environments in fault-bounded grabens⁷. The thick organic-rich shales of the Barney Creek Formation are intimately interbedded with talus breccias and sabkha-like evaporite cycles, indicating deposition in arid climates in shallow playas or near-shore lagoonal complexes¹⁷. The Yalco Formation contains thin (1–4 m) organic-rich shales (Table 1, Fig. 1b) within a series of

Received 11 February; accepted 16 May 1986.

1. Jahn, A. *Problems of the Periglacial Zone* (Polish Scientific, Warsaw, 1975).
2. Black, R. F. *Quat. Res.* 6, 3–26 (1976).
3. French, H. M. *The Periglacial Environment* (Longman, London, 1976).
4. Washburn, A. K. *Geocryology* (Wiley, New York, 1980).

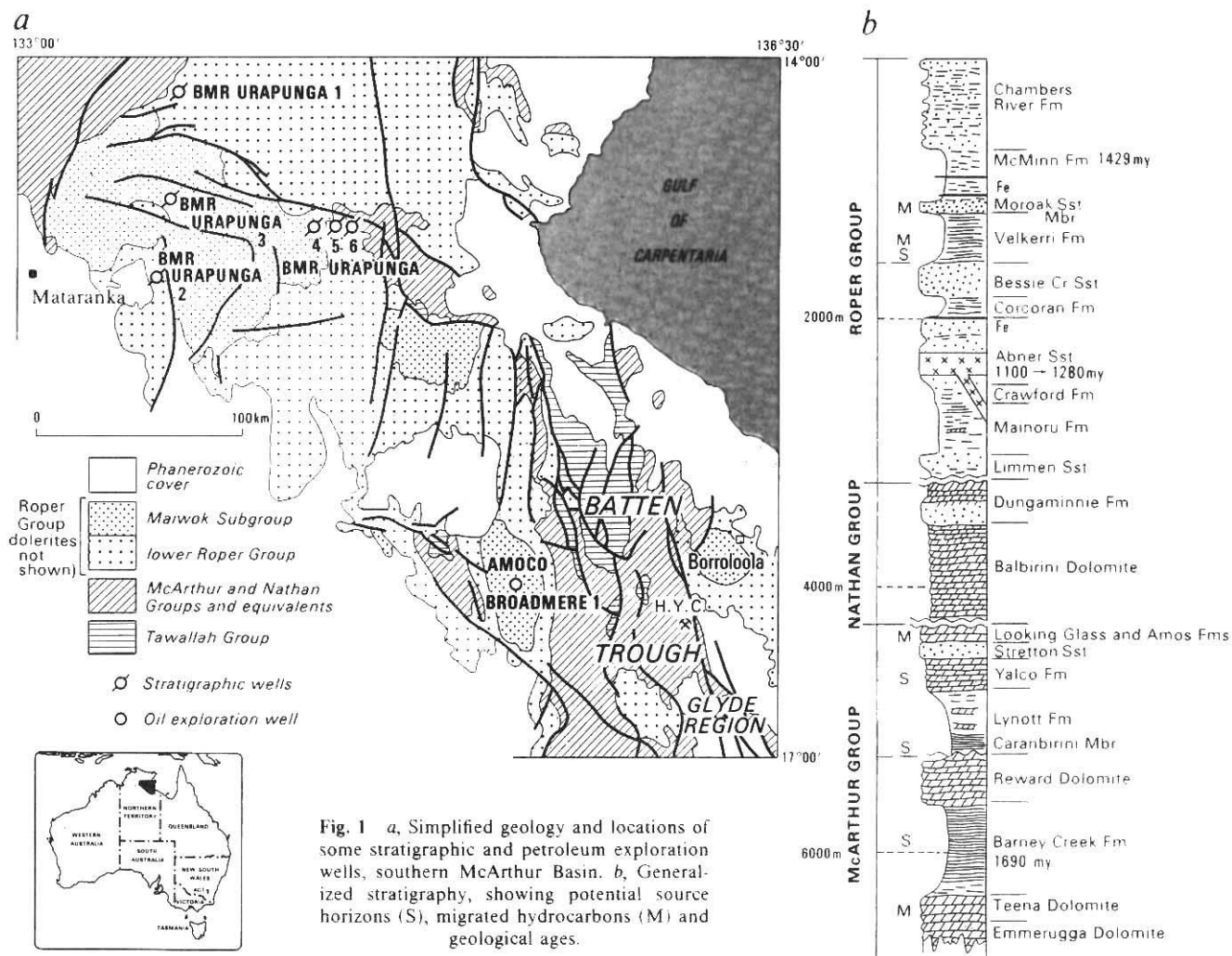


Fig. 1 a, Simplified geology and locations of some stratigraphic and petroleum exploration wells, southern McArthur Basin. b, Generalized stratigraphy, showing potential source horizons (S), migrated hydrocarbons (M) and geological ages.

variable cherty, stromatolitic carbonates and minor sandstones deposited in ephemeral lakes in a dominantly humid climate¹⁸. The Roper Group, in contrast, comprises a laterally extensive sequence of well-sorted quartz arenites and grey shales deposited in marine conditions. Depositional environments range from inner shelf for the arenites to low-energy, deeper marine environments for the organic-rich sediments of the Velkerri Formation (Table 1).

Oil bubbled from thin laminae of siltstone enclosed in black mudstone of the Velkerri Formation at a depth of 345.4 to 346.5 m in the BMR Urapunga no. 4 stratigraphic hole (Fig. 1). Preliminary GCMS (gas chromatography/mass spectrometry) analyses of the oil and marginally mature source rocks in the McArthur Basin show a preponderance of hydrocarbons derived

from organic matter of prokaryotic origin (see Fig. 2). Analyses of the C_{12+} hydrocarbon fraction from the oil revealed a mixture dominated by *n*-alkanes of low relative molecular mass without odd-over-even predominance, 1- and 2-methylalkanes, ω -cyclohexyl alkanes and a series of unresolved mixtures, possibly of monomethylalkane isomers. The last may be similar to those found in Upper Proterozoic oils from Oman¹⁹. Only a trace of pristane was detected, and sterane and hopane biomarkers were absent. In contrast, less mature Velkerri sediments contained low amounts of pristane, phytane, hopanes and steranes. Low-maturity sediments from the Barney Creek Formation contained abundant pristane, phytane, extended acyclic isoprenoids, hopanes and traces of steranes.

These results demonstrate for the first time the presence of

Table 1 Summary of source-rock characteristics of black shales, McArthur Basin

Formation (maturity)*	Organic C (%)	Potential† (kg tonne ⁻¹)	Hydrocarbons‡ (mg per g org. C)	Kerogen§ (H/C)	Organic type
McMinn (Immature)	0.7-2.9	2.3-14.6			II/III
Velkerri (Mature)	1.7-7.2	8.4-33.3	30-110	0.76-1.01	II
Yalco (Immature)	0.8-5.4	3.8-33.0	3.5-33	0.92-1.20	II
Caranbirini (Overmature)	0.2-3.4	<0.32			?
Barney Creek (Immature)	0.6-10.4	2.5-71.3	<20		I/II
(Mature)	0.8-7.6	1.3-39.6	12-87		I/II
(Overmature)	0.2-3.2	0-0.85	<6		?

* Maturity level where sampled.

† Total pyrolysis yield from Rock Eval analysis²³.

‡ Soluble hydrocarbons expressed as mg per g organic carbon²⁴.

§ From elemental analysis of isolated kerogens.

|| Interpreted organic matter type, after ref. 25.

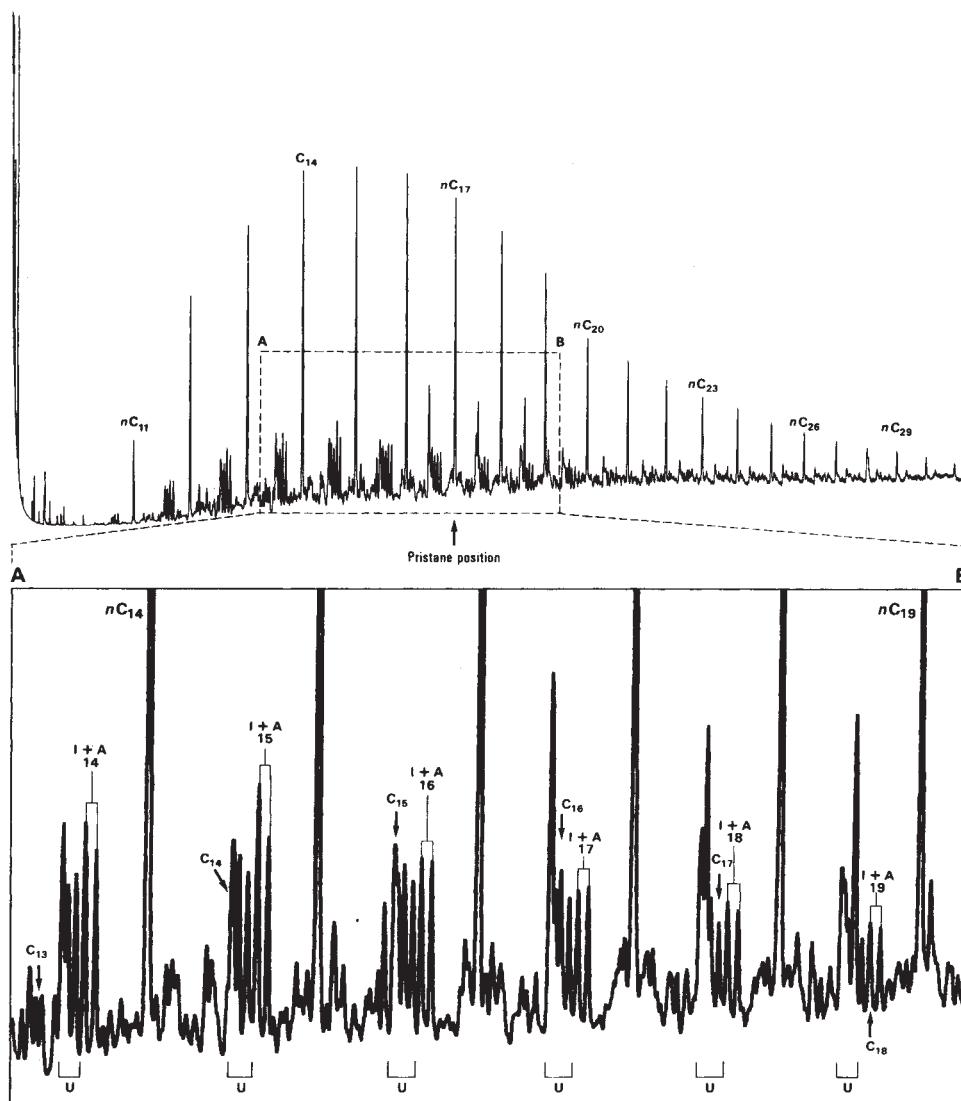


Fig. 2 Saturate fraction gas chromatogram from BMR Urupunga no. 4 oil, showing compound classes identified by GCMS. nC_{11} - nC_{29} , n -alkanes; I+A(14)-I+A(19), C_{14} - C_{19} iso- and anteisoalkanes; C_{13} - C_{18} , C_{13} - C_{18} ω -alkylcyclohexanes; U, unidentified methyl alkanes.

'live' oil and petroleum source rocks in 1,400-1,670-Myr-old sediments, and record the oldest unequivocally indigenous hydrocarbon biomarkers so far identified. They significantly change our estimate of the volume of sediments that may contain petroleum in Australia and elsewhere. Although the oil and sediments contain hydrocarbons which imply prokaryotic precursors, in general, they are not significantly different from many Phanerozoic samples, emphasizing the importance of the contribution of prokaryotic organic matter to source rocks of all ages²⁰. Subtle compositional differences in the contents of acyclic and cyclic biomarkers between the Velkerri Formation (marine) and the Barney Creek Formation (lacustrine) illustrate that environmental differences are reflected in the composition of sedimentary organic matter, even in the Proterozoic. In all the McArthur sediments, the extremely low abundance of steranes (biomarkers for eukaryotic organisms²¹) compared with hopanes, supports the concept of the rise of the eukaryotes in the latest Proterozoic (that is, after 1,400 Myr)²². However, this evidence remains equivocal until environmental influences on organic matter composition in these sediments can be determined.

We thank C. Boreham and Z. Horvath for analyses and M. R. Walter and P. J. Cook for helpful discussion. This paper is published with permission of the director of the Bureau of Mineral Resources, Geology and Geophysics. The Australian National Energy Research, Development and Demonstration Program supported this work.

Received 13 December 1985; accepted 2 June 1986.

1. Meyerhoff, A. A. *Mem. Am. Ass. Petrol. Geol.* **30**, 225-252 (1980).
2. Meyerhoff, A. A. *Dist. Lect. Ser. (Petrol. Explor. Soc. Aust.)* (August 1982).
3. Murray, G. E., Kaczov, M. J. & McArthur, R. E. *Bull. Am. Ass. Petrol. Geol.* **64**, 388-390 (1977).
4. Petrov, A. A., Mokshantsev, K. B., Fradkin, G. S. & Barykin, S. F. *Petrol. Geol.* **18**, 388-390 (1977).
5. Scott, R. W. *Wild Oil* **189**, 55-61 (1979).
6. Hobson, G. D. & Tiratsoo, E. N. *Introduction to Petroleum Geology*, 2nd edn (Gulf, Houston, 1981).
7. Jackson, M. J., Muir, M. D. & Plumb, K. A. *Bull. Bur. Miner. Resour. Geol. Geophys. Aust.* (in the press).
8. Page, R. W. *Econ. Geol.* **76**, 648-658 (1981).
9. Kralik, M. *Precamb. Res.* **18**, 157-170 (1982).
10. Muir, M. D. *Origins Life* **5**, 143-173 (1974).
11. Muir, M. D. *Alcheringa* **1**, 143-158 (1976).
12. Oehler, J. H. *Alcheringa* **1**, 315-349 (1977).
13. Oehler, D. Z. *Alcheringa* **2**, 269-309 (1978).
14. Peat, C. J., Muir, M. D., Plumb, K. A., McKirdy, D. M. & Norvick, M. J. *Bur. Miner. Resour. J. Aust. Geol. Geophys.* **3**, 1-17 (1978).
15. Muir, M. D. *Bot. J. Linn. Soc.* **86**, 1-18 (1983).
16. Muir, M. D., Armstrong, K. J. & Jackson, M. J. *Bur. Miner. Resour. J. Aust. Geol. Geophys.* **5**, 301-304 (1980).
17. Logan, R. G. & Williams, N. *7th Aust. Geol. Conv. Abstr.* **12**, 339 (Geological Society of Australia, 1984).
18. Muir, M. D., Lock, D. & von der Borch, C. C. *Spec. Publ. Soc. econ. Paleont. Miner.* **28**, 51-67 (1980).
19. Klomp, U. C. *Abstr. 12th int. Meet. org. Geochem.*, (Julich, European Association of Organic Geochemists, 1985).
20. Ourisson, G., Albrecht, P. & Rohmer, M. *Scient. Am.* **251**(2), 34-41 (1984).
21. Rohmer, M., Bouvier, P. & Ourisson, G. *Proc. natn. Acad. Sci. U.S.A.* **76**, 847-851 (1979).
22. Schopf, J. W. *Precamb. Res.* **5**, 143-173 (1977).
23. Espitalie, J., Madec, M., Tissot, B., Meening, J. J. & Leplat, P. *Offshore Tech. Conf., Prepn.* **9**, 3, 439-444 (1977).
24. Powell, T. G. *Geol. Surv. Pap. Can.* **78**-12 (1978).
25. Tissot, B., Durand, B., Espitalie, J. & Combaz, A. *Bull. Am. Ass. Petrol. Geol.* **58**, 499-506 (1974).

Amino acid epimerization in planktonic foraminifera suggests slow sedimentation rates for Alpha Ridge, Arctic Ocean

S. A. Macko & A. E. Aksu

Department of Earth Sciences, Memorial University of Newfoundland, St John's, Newfoundland, Canada A1B 3X5

Interpretations of the palaeoclimatic and palaeoceanographic histories of the Arctic Ocean^{1,2} have been hampered by conflicting chronologies obtained using different dating methods. Magnetostratigraphic and micropalaeontological data¹⁻⁴ have suggested a sedimentation rate of $\sim 1 \text{ mm kyr}^{-1}$, whereas the rate indicated by amino acid epimerization analyses on foraminifera⁵ is an order of magnitude higher. To establish a firm chronology for cores from the Alpha Ridge, Arctic Ocean, and to help resolve the controversy between the ages assigned by different dating methods, we have carried out detailed amino acid epimerization analyses on planktonic foraminifera *Neoglobobulimina pachyderma* sinistral from a long gravity core. Assuming that the temperature in the Arctic Ocean did not exceed 3°C during the past 1 Myr, the alloseucine/isoleucine (alle/Ile) ratios obtained from this core suggest a sedimentation rate of $\sim 1 \text{ mm kyr}^{-1}$, supporting previous magnetostratigraphic and micropalaeontological studies.

This study was primarily based on amino acid epimerization analyses on planktonic foraminifera extracted from core CESAR 83-102 ($85^\circ 38' \text{N}$, $111^\circ 07.1' \text{W}$), raised from a depth of 1,495 m on the Alpha Ridge, Arctic Ocean (Fig. 1). The core includes down-core sequences of sediments typical of previously studied Arctic Ocean cores (Fig. 2). Three lithostratigraphic units (K-M) identified in this core are fully discussed elsewhere⁶, and are briefly described here: lithological unit M is a partially bioturbated, foraminifera-rich, yellowish sandy to silty mud, with up to three pink-white carbonate hardground layers. Unit L consists of foraminifera-poor, light brown sandy-silty mud, and unit K is a partially bioturbated, dark brown silty-sandy mud with variable numbers of foraminifera. High-resolution seismic profiles show smooth, flat-lying reflectors over the core site⁷. Detailed examination of X-radiographs shows the absence of internal sedimentary structures usually associated with re-sedimentation. The presence of widespread bioturbational sedimentary structures suggests that the sediments recovered in this core represent pelagic or hemipelagic deposition. The magnetostratigraphic data from this core² indicate that the boundary between the Brunhes and Matuyama magnetochrons lies in unit K at $\sim 87 \text{ cm}$ depth. This boundary is generally accepted to occur at 0.73 Myr (ref. 9).

Amino acid epimerization reaction rates are species-dependent^{10,11}; therefore, only monospecific samples of *N. pachyderma* sinistral were used for the analyses. This taxon, however, shows large morphological variations and includes several morphotypes, such as *N. polusi* and *N. cryophila*, that are considered by some^{4,12} to be separate species. Therefore, only the four-chambered quadrate forms of *N. pachyderma* sinistral were used for the amino acid analyses. Twenty-seven levels were analysed in CESAR 83-102, each sample comprising a 1-cm-thick slice of the core. Despite much greater sensitivity in detection level for the amino acid technique, enabling smaller samples to be analysed, larger sample sizes were used to reduce potential laboratory contamination and to ensure sample integrity and homogeneity. About 1,000 specimens ($\sim 10 \text{ mg}$) were hand-picked from each sample; foraminiferal shells were ultrasonically cleaned in distilled water three times and each sample was rinsed with distilled water twice. Samples were sealed under nitrogen with a stoichiometrically related excess of hydrochloric acid to yield final solutions of 6 M HCl, which were then hydrolysed for 24 h at 100°C . Portions of the hydrolysate were

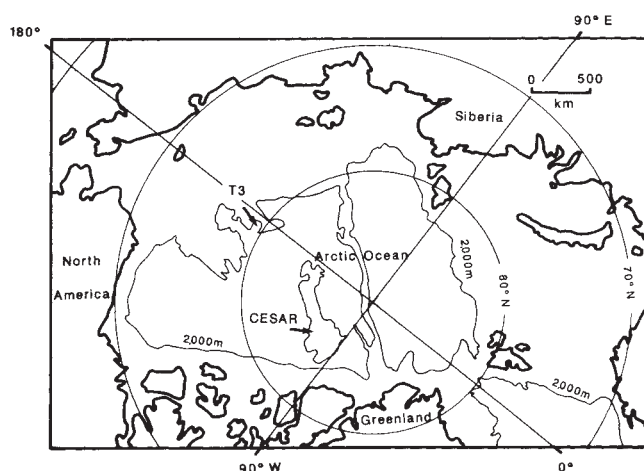


Fig. 1 Map of the Arctic Ocean showing the locations of the CESAR and T-3 ice camps.

analysed by stepwise isocratic elution with baseline separation of the alloseucine/isoleucine pair, fluorescence detection and peak integration on an ion-exchange HPLC (high-pressure liquid chromatograph) with a detection limit of $\sim 10 \text{ pM}$ (ref. 13). Triplicate analyses show that the reproducibility is ± 0.025 in the alle/Ile value. For chronostratigraphic calibration, alle/Ile was determined on two additional samples of *N. pachyderma* sinistral, extracted from well-dated ash layers 2 and 3 of the North Atlantic¹⁴ cores K708-1 and K708-7, respectively. Results of these analyses as well as an interlaboratory sample of *Saxidomus* sp. (ILC-A)¹⁵ are presented in Table 1. This

Table 1 Ratio of alle/Ile in *N. pachyderma* sinistral from Arctic Ocean core 83-102 and Atlantic Ocean cores K708-1 and K708-7, corresponding to volcanic ash layers 2 and 3 respectively¹⁴

Core	Depth (cm)	alle/Ile	Core	Depth (cm)	alle/Ile
83-102	0-1	0.034	83-102	31-32	0.253
83-102	1-2	0.041	83-102	38-39	0.235
83-102	3-4	0.075	83-102	39-40	0.288
83-102	4-5	0.091	82-102	42-43	0.223
83-102	5-6	0.110	83-102	44-45	0.220
83-102	7-8	0.085	83-102	81-82	0.280
83-102	9-10	0.119	83-102	82-83	0.267
83-102	13-14	0.204	83-102	84-85	0.287
83-102	14-15	0.193	83-102	87-88	0.322
83-102	15-16	0.140	83-102	89-90	0.318
83-102	18-19	0.209	83-102	90-91	0.287
83-102	19-20	0.214	83-102	91-92	0.275
83-102	22-23	0.229	K708-1	390	0.102
83-102	24-25	0.242	K708-7	810	0.236
83-102	27-28	0.224	ILC-A	—	0.167

ILC-A, interlaboratory calibration standard *Saxidomus* sp.¹⁵.

standard is reported for comparison as recommended by Wehmiller¹⁵.

The plot of alle/Ile in *N. pachyderma* from CESAR 83-102 against depth reveals two regions of linearity (Fig. 3). The upper 13 samples show a constant increase in alle/Ile to $\sim 23 \text{ cm}$ depth in the core, and this data set displays a highly significant correlation ($r^2 = 0.83$; Pearson r , 0.01 level). The second zone extends from ~ 23 to 93 cm depth. Linear regression over this region shows a significant correlation (0.05 level), with an r^2 value of 0.31. Sediments between 47 and 81 cm depth in the core were barren of foraminifera (Fig. 2). The alle/Ile value at the intersection of these two regions is 0.229, and is comparable to that observed elsewhere^{10,16,17}. The exact nature of the change in slope between these two regions is not fully understood. A variation in protein content (free versus bound) or the presence of more than one type of protein with different resistance to

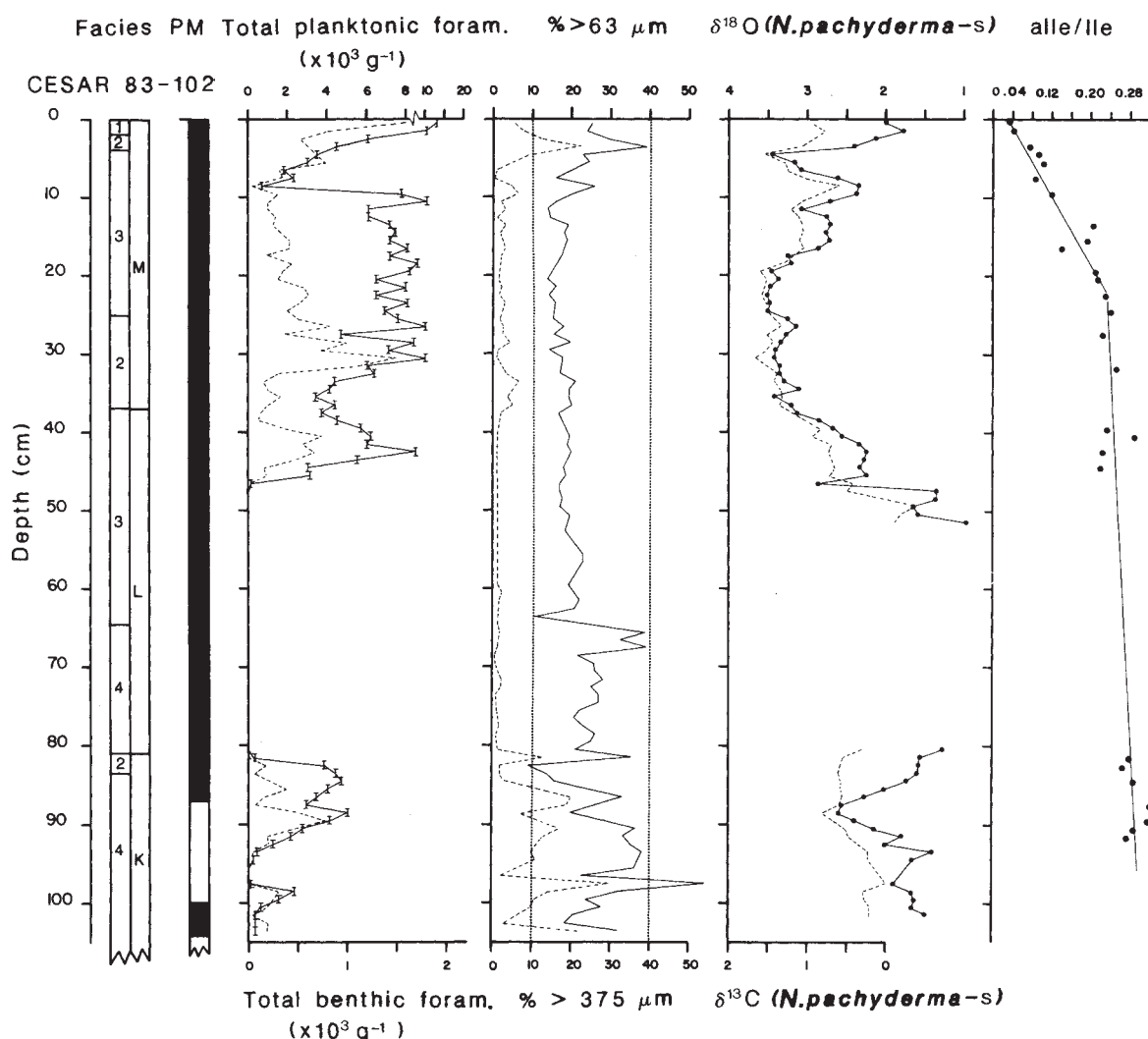


Fig. 2 Averaged amino acid epimerization data plotted against depth, together with lithofacies⁶, palaeomagnetic timescale², foraminiferal, grain size and stable isotope data⁸. Facies: 1, clayey mud; 2, carbonate hardgrounds; 3, silty mud; 4, sandy mud; units K-M are explained in text. Dashed lines refer to lower axes.

degradation or associated rates of epimerization may be the cause¹⁶.

It has been demonstrated that the epimerization reaction of isoleucine



obeys first-order reaction kinetics¹⁷⁻¹⁹. The reaction rate constants k_{ile} and k_{alle} are temperature-dependent and are governed by the Arrhenius equation:

$$k_{\text{ile}} = A e^{(-E_a/RT)} \quad (2)$$

where A is the species-dependent Arrhenius frequency factor¹¹, E_a the genus-dependent activation energy¹⁷⁻¹⁹, R the gas constant and T the temperature (K). The rate constants for the North Atlantic core samples can then be calculated using a solution to the differential rate expression²⁰:

$$\ln \left\{ \frac{1 + \frac{\text{alle}}{\text{Ile}}}{1 - K' \frac{\text{alle}}{\text{Ile}}} \right\}_t - \ln \left\{ \frac{1 + \frac{\text{alle}}{\text{Ile}}}{1 - K' \frac{\text{alle}}{\text{Ile}}} \right\}_{t=0} = (1 + K') k_{\text{ile}} t \quad (3)$$

where $K' = k_{\text{alle}} = 1/K_{\text{eq}}$, the inverse of alle/Ile at equilibrium, and t is the time. Although alloisoleucine may be present in

living foraminifera and isoleucine may generate alloisoleucine during acid hydrolysis, these contributions ($t=0$ term in (3)) are negligible^{21,22}.

The value for K_{eq} is $\sim 1.3-1.4$ (refs 17, 20, 22). Taking this value to be 1.4 (refs 20-22), with ages for ash layers 2 and 3 of 65,000 and 340,000 yr BP (ref. 14), respectively, the forward rate constants for epimerization in these two samples are calculated to be 1.551×10^{-6} and $6.803 \times 10^{-7} \text{ yr}^{-1}$, respectively, using measured alle/Ile values (Table 1). An additional rate constant of $1.545 \times 10^{-6} \text{ yr}^{-1}$ is calculated using published data⁵ from the same core, K708-7 (alle/Ile = 0.078, age = 50,000 yr BP, depth = 140 cm). Ages calculated using first-order reversible kinetics of epimerization in marine fossils are found to be optimal for samples with alle/Ile ≤ 0.25 (refs 16, 20-23); because the value in ash layer 3 is near this limit, the alle/Ile and derived rate constant for this sample are not used for further calculations. The Arrhenius frequency factor ($\log A$) is calculated to be 15.70 using the average of the remaining two rate constants and assuming an activation energy of $113.4 \text{ kJ mol}^{-1}$, which was calculated from a bulk foraminiferal sample¹⁷. This value is nearly identical to the value of 15.77 calculated²⁰ from a North Atlantic core. From these calculations, the relationship between forward epimerization rate constant and temperature for *N. pachyderma* can be expressed as:

$$\log k_{\text{ile}} \text{ yr}^{-1} = 15.70 - 5,939/T \quad (4)$$

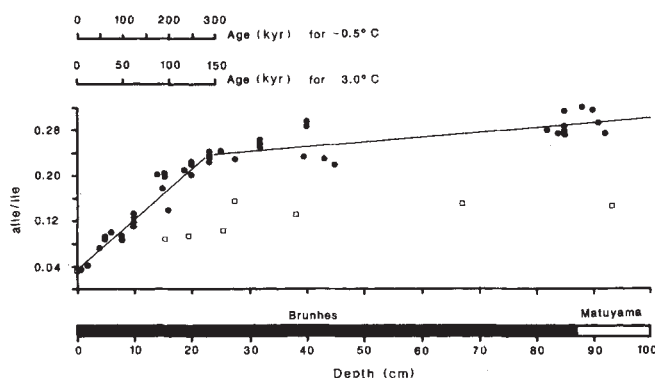


Fig. 3 Epimerization of isoleucine in *N. pachyderma* sinistral versus depth in core 83-102. Age scale discussed in text, palaeomagnetic scale from ref. 2. □, from ref. 5, re-plotted based on the Brunhes/Matuyama boundary as the reference point.

Present-day Arctic water temperatures vary from -0.5°C in the Canada Basin and the Alpha Ridge bottom waters to a maximum in the water column not exceeding $+3^{\circ}\text{C}$ (ref. 24). First-order epimerization rate constants based on the above temperature range are 8.194×10^{-7} and 1.548×10^{-6} for -0.5 and 3°C , respectively. Relative ages of sediment can then be calculated using equation (2) (Fig. 3). For example, the sample at 20 cm depth, with an alle/Ile of 0.214 (average of 3 analyses) yields ages of 256,100 and 135,600 yr BP for the maximum- and minimum-temperature rate constants. These ages give sedimentation rates of $0.8\text{--}1.5\text{ mm kyr}^{-1}$ for the Alpha Ridge region of the Arctic Ocean.

This deposition rate is consistent with previous magnetostratigraphic and biostratigraphic studies^{1-4,8,12,24-26}, but conflicts strongly with the conclusions of the recent amino acid epimerization study⁵ of the Arctic Ocean core T3-67-11. These authors⁵ have argued that a significant portion of the foraminifera analysed from their core were reworked, and that the alle/Ile measured on the deepest sample (260 cm) was closest to the *in situ* ratio in this core. Consequently, their age estimate, using calculations similar to those presented here, was based solely on this sample. The rate constant derivation for their estimate was based on benthic foraminifera *C. lobatulus*. There is no evidence for significant reworking of fauna in studied CESAR cores^{2,6,8}. Palynological data² from a long CESAR core show several first-appearance and last-appearance data of dinoflagellate and pollen and spore taxa in the appropriate sequence, generally supporting little reworking of fauna and flora. It has been demonstrated^{6,24} that there is a strongly correlatable down-core sequence of 13 lithological units in >500 cores collected from the Arctic Ocean. Several studies have indicated that there are also correlatable faunal and floral abundance and assemblage variations associated with the lithostratigraphic units²⁻⁴. The palaeomagnetic data from several cores from the Arctic Ocean have shown the presence of normal and reversed polarity chrons that are correlated with known magnetic timescales^{1,2,25,26}. Core T3-67-11, studied by Sejrup *et al.*⁵, shows litho-, bio- and magnetostratigraphic variations similar to other cores from the Arctic Ocean^{1,4}. Therefore, the assumption that most foraminifera from core T3-67-11 are reworked may not be warranted.

The foraminifera analysed from core T3-67-11 (ref. 5) show a down-core increase in alle/Ile to a depth of 120 cm. This trend is similar to that observed in CESAR core 83-102 (Fig. 3). The reason for the difference in the slopes of the two data sets is not clear. A variation of this magnitude requires a decrease of $\sim 3\text{--}5^{\circ}\text{C}$ in the average temperature of epimerization for T3 with respect to CESAR sites. Assuming that foraminifera recovered from this core were not reworked and that the average bottom water temperature was -0.5°C , the alle/Ile values reported by Sejrup *et al.*⁵ yield respective sedimentation rates of ~ 1.9 and 2.6 mm kyr^{-1} for samples from 33 and 35 cm depth in T3-67-11. These rates are only slightly higher than the value of 1.6 mm kyr^{-1} predicted by the palaeomagnetic data from this core¹. We

recognize that a single sedimentation rate for the entire Arctic Ocean is unlikely. Based on the above calculations, the sedimentation rates in the Canada Basin (T3) are slightly higher than those calculated for the Alpha Ridge (CESAR); however, these rates are 10–15 times slower than that suggested by Sejrup *et al.*⁵.

The amino acid data presented in this paper independently demonstrate a slow sedimentation rate ($\sim 1\text{ mm kyr}^{-1}$) for the Alpha Ridge region of the Arctic Ocean, supporting previous magneto- and biostratigraphic timescales. Through detailed multidisciplinary investigations, the chronology of the Arctic Ocean sediments is believed to have been resolved. A reliable chronology is an important prerequisite in studying the palaeoclimatic and palaeoceanographic histories of the Arctic Ocean and its role in Northern Hemisphere glacial/interglacial cycles.

We thank M. J. Keen and H. R. Jackson of the Bedford Institute of Oceanography for making the CESAR 83-102 samples available and K. Pulchan for assistance in laboratory measurements. This study was supported by the Natural Sciences and Engineering Research Council of Canada. Acknowledgement is made to the donors of The Petroleum Research Fund, administered by the American Chemical Society, for partial support. W. F. Ruddiman provided North Atlantic samples from cores K708-1 and K708-7 and J. F. Wehmiller supplied the interlaboratory standard ILC-A. We thank J. F. Wehmiller, J. A. Welhan, D. J. W. Piper and C. Pereira for critical comments.

Received 10 February; accepted 2 May 1986.

- Hunkins, K., Be, A. W. H., Opdyke, N. D. & Mathieu, G. *The Late Cenozoic Glacial Ages* (ed. Turekian, K. K.) 215–237 (Yale University Press, 1971).
- Aksu, A. E. & Mudie, P. J. *Nature* **318**, 280–283 (1985).
- Clark, D. L. *Bull. geol. Soc. Am.* **82**, 3313–3323 (1971).
- Herman, Y. *Marine Geology and Oceanography of the Arctic Seas* (ed. Herman, Y.) 283–348 (Springer, Berlin, 1974).
- Sejrup, H. P., Miller, G. H., Brigham-Grette, J., Lovlie, R. & Hopkins, D. *Nature* **310**, 772–775 (1984).
- Mudie, P. J. & Blasco, S. M. *Geol. Surv. Pap. Can.* **84-22**, 59–99 (1985).
- Jackson, H. R. *Geol. Surv. Pap. Can.* **84-22**, 19–23 (1985).
- Aksu, A. E. *Geol. Surv. Pap. Can.* **84-22**, 115–124 (1985).
- Harland, W. B. *et al.* *A Geologic Time Scale* (Cambridge University Press, 1982).
- King, K. & Neville, C. *Science* **195**, 1333–1335 (1977).
- Miller, G. H., Sejrup, H. P., Mangerud, J. & Anderson, B. G. *Boreas* **12**, 107–124 (1983).
- Herman, Y. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **6**, 251–276 (1969).
- Hare, P. E., St. John, P. A. & Engel, M. H. *Chemistry and Biochemistry of Amino Acids* (ed. Barnett, G. C.) 415–425 (Chapman and Hall, London, 1984).
- Ruddiman, W. F. & McIntyre, A. *Soc. Am. Mem.* **145**, 111–146 (1976).
- Wehmiller, J. F. *Quat. Res.* **22**, 109–120 (1984).
- Schroeder, R. A. & Bada, J. L. *Earth Sci. Rev.* **12**, 347–391 (1976).
- Bada, J. L. & Schroeder, R. A. *Naturwissenschaften* **62**, 71–79 (1975).
- Miller, G. H. & Hare, P. E. *Biogeochemistry of Amino Acids* (eds Hare, P. E., Hoering, T. C. & King, K.) 415–444 (New York, 1980).
- Mitterer, R. M. *Geology* **2**, 425–428 (1975).
- Bada, J. L. & Schroeder, R. A. *Earth planet. Sci. Lett.* **15**, 1–11 (1972).
- Bada, J. L. & Mann, E. H. *Earth Sci. Rev.* **16**, 21–55 (1980).
- Wehmiller, J. F. & Hare, P. E. *Science* **173**, 907–911 (1971).
- King, K. & Hare, P. E. *Yb. Carnegie Instn. Wash.* **71**, 596–598 (1972).
- Coachman, K. L. & Aagaard, K. *Marine Geology and Oceanography of the Arctic Seas* (ed. Herman, Y.) 1–72 (Springer, Berlin, 1974).
- Clark, D. L., Whitman, R. R., Morgan, K. A. & Mackey, S. D. *Geol. Soc. Am. spec. Pap.* **181** (1980).
- Steuerwald, B. A., Clark, D. L. & Andrew, J. T. *Earth planet. Sci. Lett.* **5**, 79–85 (1968).
- Clark, D. L. *Bull. geol. Soc. Am.* **81**, 3129–3134 (1970).

Thermal and mechanical constraints on the lithosphere beneath the Marquesas swell

Karen M. Fischer*†, Marcia K. McNutt* & Loren Shure†

* Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

† Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

The Marquesas Islands, like most active or recently active hotspot chains, lie on anomalously shallow sea floor, and their subsidence history points to a fundamentally thermal origin for the swell¹, as do heat flow and subsidence data for the swells beneath Bermuda² and Hawaii^{3,4}. Here we apply a new linear filtering technique to geoid height and depth anomaly data over the Marquesas, to constrain the depth to the high-temperature, low-density material responsible for elevating the swell. We determine an isostatic compensation depth of 45 ± 5 km for the swell, and best-fitting values for lithospheric flexural rigidity and island load density of 5×10^{22} N m and $2,700 \text{ kg m}^{-3}$, respectively. The swell's shallow compensation depth and low flexural rigidity call for anomalously high temperatures within the lithosphere^{5,6}.

Crough and Jarrard¹ have computed an isostatic root depth of 30 ± 40 km for the Marquesas swell using a geoid height–depth anomaly transfer function. Their result suggests anomalously high temperatures well into the lithosphere, as isostatic compensation depth provides a fairly direct estimate of the depth to the thermal anomaly². Their study is compatible with the low flexural rigidity of 4×10^{22} N m for the Marquesas derived by Cazenave *et al.*⁷, given that the thickness of the elastic plate supporting intraplate volcanoes has been shown to be thermally

controlled⁸. However, the transfer function method, and other current techniques for estimating swell compensation depth, use assumed values for swell height and width taken from geoid or depth anomaly data. In the Marquesas, the overlap in wavelength between the isostatically compensated swell and the mechanically supported islands makes it difficult to obtain estimates of swell height that are unbiased by island-related topography.

In the present study we apply a new method for estimating swell compensation depth to updated depth anomaly and geoid height maps over the Marquesas. The technique, described more fully in ref. 9, does not require initial estimates of swell amplitude or wavelength, nor does it require that the swell and islands be separated in the spectral domain.

The two primary data sets discussed here are seafloor depth anomalies and geoid height anomalies computed between 160 and 125° W, and 5 and 30° S, an area loosely covering French Polynesia. As we are interested only in hotspot signatures, we attempt to remove from the data all effects of non-hotspot-related processes.

Bathymetry at 5' intervals from the Synbaps data set¹⁰ were averaged to a new spacing of 15' and then corrected for sound velocity variations with a quadratic fit to the Matthews tables¹¹ (L. Wixom, personal communication). The effect of sediment loading was removed according to the empirical relation given in ref. 12, using sediment thicknesses measured from ref. 13. Finally, we accounted for thermal subsidence due to the cooling of the oceanic lithosphere using the depth–age relationship for the North Pacific from ref. 14. Seafloor ages were inferred from the isochron locations of refs 1, 15 and H. W. Menard (unpublished data), and the fracture zone identifications of ref. 16. We estimate the maximum errors in the depth anomaly map for the Marquesas (Fig. 1a) to be less than the 250-m contour interval.

Geoid height data are from the GEOS-3 and Seasat altimeter missions as compiled in ref. 17, at a 15' spacing. Components of the field with wavelengths $>4,000$ km were eliminated by subtracting the observed field up to degree 10 using the Gem-9

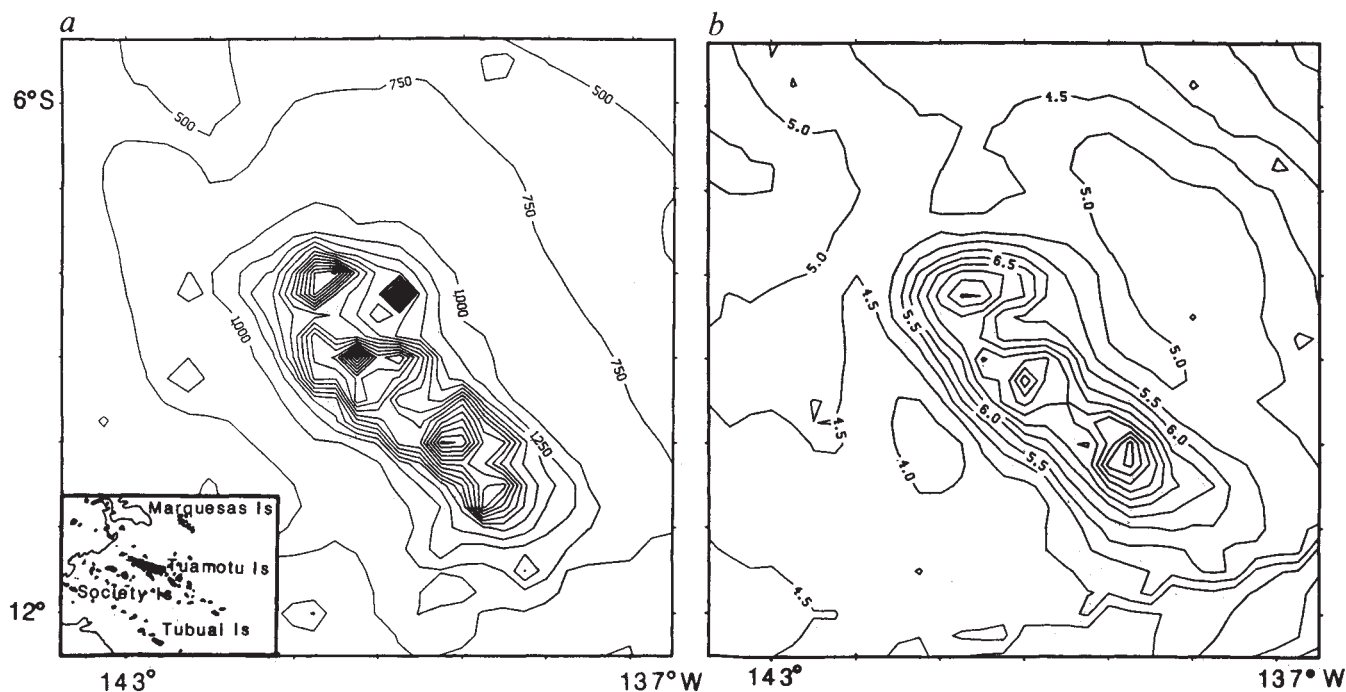


Fig. 1 Maps of reduced depth (a) and geoid (b) anomaly over the Marquesas Islands. a, Contour interval 250 m. The inset shows the major bathymetric features in the area of French Polynesia between 160 and 125° W, and 5 and 30° S (after ref. 1). The single contour in the inset is at 5 km depth. b, Contour interval 0.5 m. Note the flexural moat signatures to the north-east and south-west of the islands.

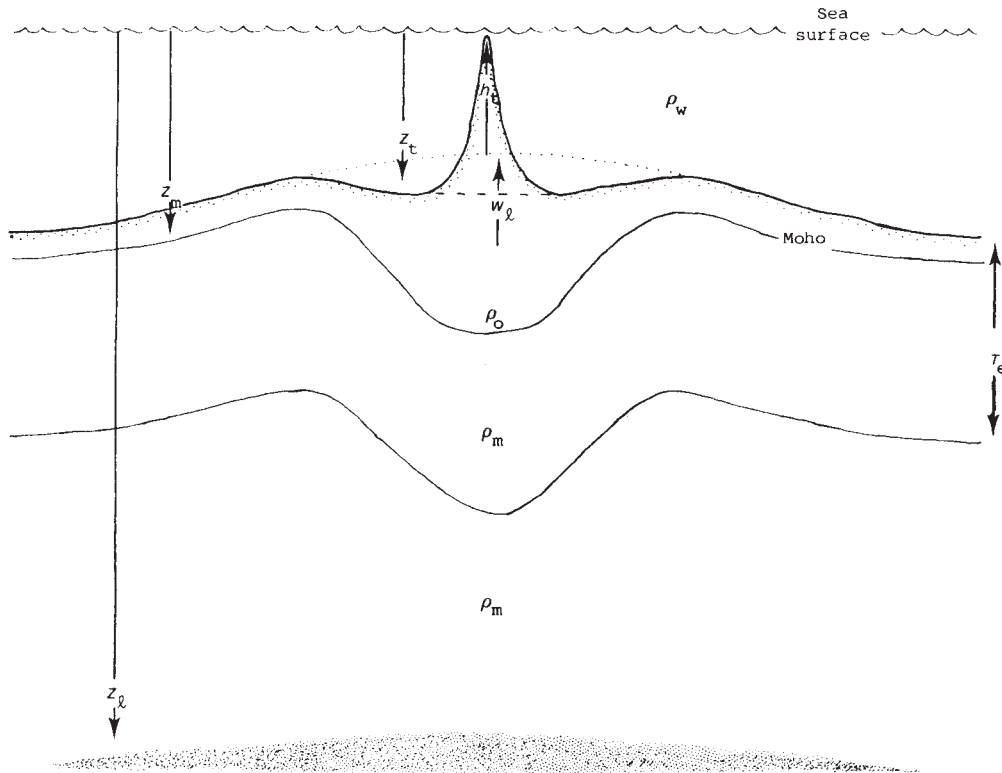


Fig. 2 Schematic cross-section of a swell (from ref. 9). The densities of sea water, volcanic rock and the mantle are ρ_w , ρ_0 and ρ_m , respectively. The height of the surface load is h_t , superimposed on the height of the swell, w_l . Average depth to the sea floor, Moho and swell compensation (shaded, density ρ_l) are z_t , z_m and z_l , respectively. T_e is the thickness of the elastic plate.

coefficients¹⁸ for the spherical harmonic expansion of the Earth's geoid, and the data were then corrected for the decrease in geoid height due to thermal cooling with age using the geoid height-age relationship of ref. 19 and the North Pacific thermal parameters of ref. 14. At this point the residual geoid showed a long-wavelength NE-SW trend over the Marquesas, in addition to the swell and island signatures. This feature was removed by subtracting the best-fitting plane (in the least-squares sense) from the data between 144 and 136.5°W, and 5 and 12.5°S, leaving the residual geoid shown in Fig. 1b.

Our method for estimating the compensation depth of intra-plate swells, explained in more detail in ref. 10, assumes an elastically compensated island topography, h_t , superimposed on a swell with amplitude w_l above normal seafloor depth, due to a sub-plate load with thickness h_l at an average depth z_l , which is the effective compensation depth of the swell (Fig. 2). The observables are the total age-corrected geoid anomaly, N , and the total age- and sediment-corrected depth anomaly, $h = h_t + w_l$. If we were to make the false assumption that the observed geoid is due only to the compensation of the observed topography at the Moho by elastic flexure of the oceanic crust, we would predict a geoid anomaly $\hat{N}$, which we define as

$$\hat{N}(k_x, k_y) = \frac{G(\rho_0 - \rho_w)}{gk} \left\{ e^{-2\pi k z_t} - \frac{e^{-2\pi k z_m}}{1 + k^4 \alpha} \right\} (H_t + W_l) \quad (1)$$

$$= F_1(k) \cdot H(k_x, k_y)$$

where $\alpha = (2\pi)^4 D / ((\rho_m - \rho_0)g)$; $D = ET_e^3 / 12(1 - \nu^2)$; Young's modulus $E = 8 \times 10^{10} \text{ N m}^{-2}$; T_e is the elastic plate thickness; Poisson's ratio $\nu = 0.25$; k_x and k_y are the horizontal wavenumbers; $k = (k_x^2 + k_y^2)^{1/2}$; ρ_0 , ρ_w and ρ_m are the densities of volcanic rock, water and mantle, respectively; G is the gravitational constant; g is the acceleration due to gravity; z_t is the average depth of topography; z_m is the average depth of the Moho; and upper-case letters are the Fourier transforms of their lower-case counterparts.

By subtracting $\hat{N}$ from the observable N , we obtain a form of reduced geoid anomaly, N_R , which in turn enables us to infer

the swell topography, W_l :

$$W_l(k_x, k_y) = gk / G \left\{ \left(\frac{\rho_0 - \rho_w}{1 + k^4 \alpha} + (\rho_m - \rho_0) \right) e^{-2\pi k z_m} - ((\rho_m - \rho_w) + (\rho_m - \rho_0) k^4 \alpha) e^{-2\pi k z_l} \right\}^{-1} N_R(k_x, k_y) \quad (2)$$

$$= F_2(k, z_l) \cdot N_R(k_x, k_y)$$

The filter F_2 corrects for the compensation falsely assumed in $\hat{N}$ and N_R , and provides an estimate of swell amplitude independent of island topography.

Our method, then, is to assign values for D , ρ_0 , ρ_m , z_t and z_m , the geophysical parameters in the filter F_1 which describe the geoid signature due to plate flexure. We apply the filter F_1 to observed depth anomaly data to obtain $\hat{N}$, and then we subtract $\hat{N}$ from the total observed geoid N , which gives us N_R . We then assume a value of z_l , apply the filter F_2 to N_R , and obtain an estimate of the swell topography. The best-fitting z_l values are those which generate profiles of predicted swell topography that match the swell outline seen in the total observed topography. We then test a given value of z_l in two dimensions by subtracting the predicted swell topography from the observed depth anomalies. The remaining topography should show only the volcanic islands and the signature of their flexural compensation.

Values of z_l between 25 and 55 km were tested in 5-km increments, along with many values of ρ_0 and D . Load densities between 2,500 and 2,800 kg m^{-3} were considered to be reasonable^{20,21}, although the chemistry of the Marquesas²² suggests that values in the upper end of this range are the most plausible. Values of flexural rigidity were varied from the low value of $2 \times 10^{22} \text{ N m}$ suggested by McNutt and Menard²⁰ for the Cook and Tuamotu islands, up to 10^{23} N m . In all cases ρ_m was set to 3,300 kg m^{-3} , z_m to 11 km and z_t to 6 km.

The parameter values that produce the best fit to the observed swell (Fig. 3a) and yield a residual topography with clear island signatures (Fig. 4) are $z_l = 45 \pm 5 \text{ km}$, $D = 5 \times 10^{22} \text{ N m}$ and $\rho_0 =$

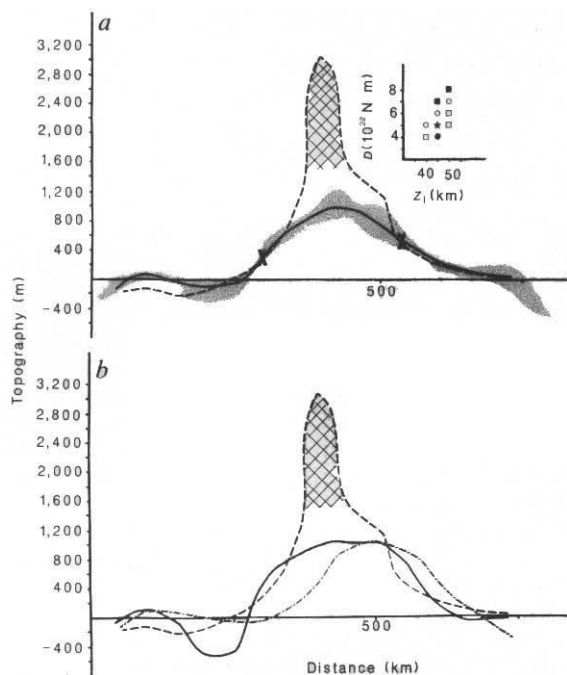


Fig. 3 *a*, Profiles of observed topography (dashed line) and swell predicted by $z_1 = 45$ km, $D = 5 \times 10^{22}$ N m and $\rho_0 = 2,700$ kg m $^{-3}$ (solid line). The cross-hatched peak gives an example of a typical island feature superimposed on the swell, and the shading indicates the profiles of the envelope of plausible parameter values, shown in the inset ($\rho_0 = \bullet, 2,800$; $\square, 2,700$; $\circ, 2,600$; $\blacksquare, 2,500$ kg m $^{-3}$; $\star$, best-fitting parameters). The positions of the flexural moats seen in the observed geoid are shown by crosses (see also Fig. 4). Note that the observed topography here and in *b* has been shifted downwards by a constant value of 500 m. *b*, Profiles of two models whose predicted swells do not fit the observed swell topography (dashed line) as well as do the models in *a*. Solid line, $z_1 = 50$, $D = 5 \times 10^{22}$ N m, $\rho_0 = 2,600$ kg m $^{-3}$; dash-dot line, $z_1 = 40$ km, $D = 5 \times 10^{22}$ N m, $\rho_0 = 2,700$ kg m $^{-3}$.

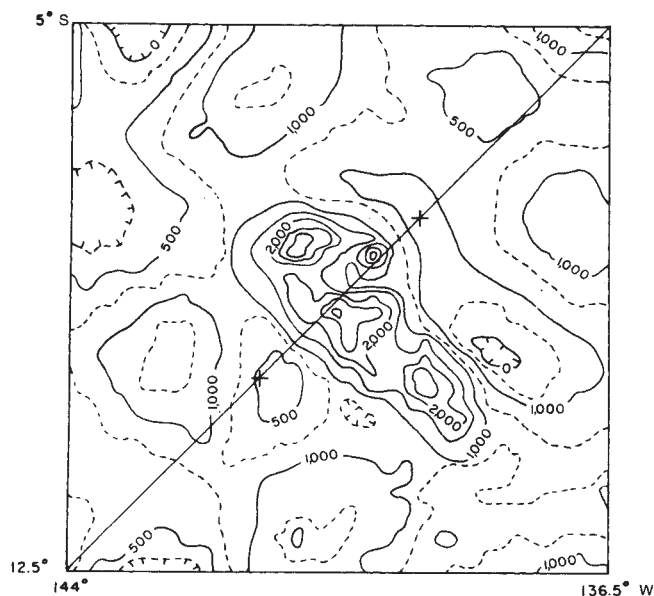


Fig. 4 Residual topography over the Marquesas Islands produced by the preferred values, $z_1 = 45$ km, $D = 5 \times 10^{22}$ N m and $\rho_0 = 2,700$ kg m $^{-3}$. Note the signatures of the volcanic islands and their flexural compensation. The diagonal line shows the locations of the profiles in Fig. 3, and the crosses correspond to the moat positions seen in the observed geoid. Solid contours are at 500-m intervals, dashed contours at 250-m intervals.

2,700 kg m $^{-3}$. The diagonal line in Fig. 4 indicates the location of the profiles in Fig. 3. The only assumption made in fitting predicted swell profiles to observed topography is that off-centre, shorter-wavelength features, such as the dotted area in Fig. 3*a*, are mechanically compensated islands. Given the ambiguity in separating island and swell contributions near the crest of the swell, preference was given to those models which best matched the flanks of the observed topography, plus or minus a constant value. The approximate locations of the flexural moats were obtained from the observed geoid (Fig. 1*b*), and allowance was made for their signature in fitting the predicted swell to observed topography profiles. Bathymetric highs corresponding to island locations, a clear flexural moat and an absence of broad highs or lows corresponding to the swell were the features sought for in the residual topography maps.

Surrounding the preferred model is an envelope of parameter values which produce less adequate fits to the swell profile (shown in the shaded area in Fig. 3*a*) and similar residual topography patterns. The range of acceptable flexural rigidity values within this envelope (shown in Fig. 3*a* inset) represents the error bars on the best-fitting flexural rigidity value of 5×10^{22} N m, and is largely constrained by the 15' spacing of the data values. An examination of the parameter envelope suggests that for a given surface load density, an increase in compensation depth may be roughly balanced by an increase in flexural rigidity. Alternatively, increasing z_1 or D offsets a decrease in ρ_0 . For instance, for the value of $\rho_0 = 2,800$ kg m $^{-3}$ used by Cazenave *et al.*⁷ and our preferred compensation depth of 45 km, we obtain their flexural rigidity of 4×10^{22} N m. For comparison, Fig. 3*b* shows two models which lie outside the envelope of best-fitting parameters.

We also checked the fit of predicted swell heights to observed topography for profiles to the north-west and south-east of those shown in Fig. 3. Over the islands to the north we obtain an identical range of best-fitting parameters. To the south, however, models with deeper compensation depths and larger flexural rigidities are called for both by swell profiles and by the apparent widening of the island moat seen in Fig. 4. Given the younger ages of the islands in this area²³, heat from a lithospheric source would have had less time to elevate lithospheric isotherms conductively, resulting in a thicker apparent lithosphere than is observed in the north.

A compensation depth of 45 km certainly suggests the presence of anomalously buoyant material within the lithosphere. If we equate the geoid anomaly due to an infinitesimally thin layer at $z_1 = 45$ km with the geoid anomaly due to a uniform column of low-density material⁹, the upper boundary of the low-density mass extends well into the upper half of the lithosphere, perhaps to within several kilometres of the surface, depending on the assumed thickness of the column. The actual depth distribution of the low-density material is, however, notably non-unique, and so we look to thermal modelling to shed some light on the significance of z_1 . Some of the Marquesas swell signature may be due to chemically defined low-density material, but as subsidence data¹ and low flexural rigidities suggest that thermal anomalies probably dominate, here we will address only possible thermal models.

Timescales associated with pure conduction are too long to explain the Marquesas swell. Assuming a lithospheric thickness of 70 km, ~ 20 Myr (at least twice the age of the oldest Marquesas volcanism²³) is required for a thermal anomaly to propagate from the base of the lithosphere to at least 45 km depth. Furthermore, if the base of the conducting region is allowed to migrate upwards as the lithosphere warms, the excess mantle heat flux must be at least four times normal in order to obtain plate thinning rates consistent with the age of the Marquesas²⁴.

Thermal structures produced by simple models of large-scale convection beneath the lithosphere are also incapable of explaining the shallow compensation depth. Calculations involving convection of a constant-viscosity material beneath a conducting

lithosphere of normal thickness (70 km) produce apparent compensation below the base of the plate²⁵, too deep to explain the observed value of 45 km, and decreases in the thickness of the conducting layer will not sufficiently diminish the calculated compensation depths²⁵. However, Robinson *et al.*²⁶ suggest that convection within certain non-uniform viscosity structures may produce thermal anomalies below 70 km depth which merely mimic compensation at depths shallower than the conducting lid thickness. More complex viscosity models may produce an apparent compensation depth at 45 km without the necessity of invoking a thinner conducting lid beneath the Marquesas.

Alternatively, Fleitout *et al.*²⁷ have proposed that if one accounts for both the temperature and pressure dependence of viscosity, a mantle heat source is liable to create convective instabilities which will erode the base of the lithosphere. This process may continue until the anomalously warm material encounters an abrupt decrease in lithospheric density, such as at the base of the basalt-depleted zone at ~40 km depth²⁸. According to this model, the 45-km compensation indicates that the lithosphere beneath the Marquesas Islands has been thinned to the maximum possible extent.

Using a linear filtering technique which requires a minimum of assumptions about swell height and wavelength⁹ we obtain a compensation depth for the Marquesas swell of 45 ± 5 km, a density of $2,700 \text{ kg m}^{-3}$ for the volcanic islands, and a flexural rigidity of $5 \times 10^{22} \text{ Nm}$ for the lithosphere. While our flexural rigidity value is comparable to that of Cazenave *et al.*⁷, our compensation depth is somewhat deeper and more tightly constrained than the results of Crough and Jarrard¹. These low values for compensation depth and flexural rigidity call for anomalously high temperatures within the lithosphere and can be used to constrain models for the formation of hotspot swells.

We thank Jim Marsh for providing the geoid data, Linda Meinke for assistance with preparing the geoid and bathymetry databases, and Barry Parsons, Steve Daly, Luce Fleitout and Beth Robinson for helpful comments. H. W. Menard provided previously unpublished isochron identifications. This project was funded by NSF OCE-8409157 and ONR N00014-82-C-0019. Woods Hole contribution 6089.

Received 2 December 1985; accepted 30 April 1986.

An enantiornithine bird from the Lower Cretaceous of Queensland, Australia

R. E. Molnar

Queensland Museum, Queensland Cultural Centre, South Brisbane, Queensland 4101, Australia

Current classification of birds recognizes four subclasses: the Archaeornithes, the Enantiornithes, the Odontornithes and the Neornithes. Three of these subclasses were proposed during the nineteenth century, but the Enantiornithes was proposed only five years ago, based on material from the Lecho Formation (Maastriichtian) of Salta Province in Argentina¹. Further material from Asia (Mongolia) and North America (Mexico) has been referred to the Enantiornithes², all of it from the Campanian. Now, in addition, a small tibiotarsus found in the Lower Cretaceous (Albian) of Queensland not only represents the first bony material of Mesozoic birds from Australia, but is also the first reported occurrence of an enantiornithine bird in the Lower Cretaceous.

The unique structure of the distal articulation of the enantiornithine tibiotarsus, with a transversely elongate medial condyle and a markedly narrower lateral condyle, identifies this tibiotarsus as enantiornithine. The specimen (QM F12992) was recovered from a limestone flag of the marine Toolebuc Formation. It was found on Warra Station, on the east side of the Hamilton River, near Hamilton Hotel, west Queensland. Further details of the stratigraphy and locality are given elsewhere³. The tibiotarsus was one of several small tetrapod elements recovered by acetic acid solution of the limestone and inspection of the residue. It was associated with plentiful small teleost remains as well as remains of other fish, turtles, ichthyosaurs and a pterosaur³.

The tibiotarsus indicates a bird the size of the European blackbird *Turdus merula*, or the American robin *Turdus migratorius*, and shows distinctions from the Salta enantiornithine material which render it desirable to erect a new genus

- Crough, S. T. & Jarrard, R. D. *J. geophys. Res.* **86**, 11763-11771 (1981).
- Detrick, R. S., Von Herzen, R. P., Parsons, B., Sandwell, D. & Dougherty, M. *J. geophys. Res.* **91**, 3701-3723 (1986).
- Detrick, R. S., Von Herzen, R. P., Crough, S. T., Epp, D. & Fehn, U. *Nature* **292**, 142-143 (1981).
- Von Herzen, R. P., Detrick, R. S., Crough, S. T., Epp, D. & Fehn, U. *J. geophys. Res.* **87**, 6711-6723 (1982).
- Menard, H. W. & McNutt, M. K. *J. geophys. Res.* **87**, 8570-8580 (1982).
- McNutt, M. K. *J. geophys. Res.* **89**, 11180-11194 (1984).
- Cazenave, A., Lago, B., Dominh, K. & Lambeck, K. *Geophys. J. R. astr. Soc.* **63**, 233-252 (1980).
- Watts, A. B., Bodline, J. & Steckler, M. *J. geophys. Res.* **85**, 6369-6376 (1980).
- McNutt, M. & Shure, L. *J. geophys. Res.* (in the press).
- Van Wykhouse, R., SYNBAPS (Synthetic Bathymetric Profiling Systems), *Tech. Rep. TR-233* (Naval Oceanographic Office, Washington, D.C., 1973).
- Matthews, D. J. *Tables of the Velocity of Sound in Pure Water and Sea Water* (HMSO, London, 1939).
- Le Douran, S. & Parsons, B. *J. geophys. Res.* **87**, 4715-4722 (1982).
- Ludwig, W. J. & Houtz, R. E. *Isopach Map of Sediments in the Pacific Ocean Basin and Marginal Sea Basins* (American Association of Petroleum Geologists, Tulsa, 1979).
- Parsons, B. & Sclater, J. G. *J. geophys. Res.* **82**, 802-827 (1977).
- Sclater, J. G., Parsons, B. & Jaupart, C. *J. geophys. Res.* **86**, 11535-11552 (1981).
- Mammerickx, J., Smith, S., Taylor, I. & Chase, T. *Topography of the South Pacific* (Scripps Institution of Oceanography, La Jolla, 1975).
- Marsh, J. G., Brenner, A. C., Beckett, B. D. & Martin, T. V. *J. geophys. Res.* **91**, 3501-3506 (1986).
- Lerch, F. J., Klosko, S. M., Laubscher, R. E. & Wagner, C. A. *J. geophys. Res.* **84**, 3897-3916 (1979).
- Parsons, B. & Richter, F. M. *Earth planet. Sci. Lett.* **51**, 440-450 (1980).
- McNutt, M. K. & Menard, H. W. *J. geophys. Res.* **83**, 1206-1212 (1978).
- Lambeck, K. *Geophys. J. R. astr. Soc.* **67**, 91-114 (1981).
- Barszczus, H. G. & Liotard, J.-M. *C. r. heb. Séanc. Acad. Sci., Paris* **299**, 61-64 (1984).
- Duncan, R. A. & MacDougall, I. *Earth planet. Sci. Lett.* **21**, 414-420 (1974).
- Spohn, T. & Schubert, G. *Tectonophysics* **94**, 67-90 (1983).
- Parsons, B. & Daly, S. *J. geophys. Res.* **88**, 1129-1144 (1983).
- Robinson, E. M., Parsons, B. & Daly, S. *Earth planet. Sci. Lett.* (submitted).
- Fleitout, L., Yuen, D. & Froidevaux, C. *Annls Geophysicae* (submitted).
- Oxburgh, E. R. & Parmentier, E. M. *J. geol. Soc. Lond.* **133**, 343-355 (1977).

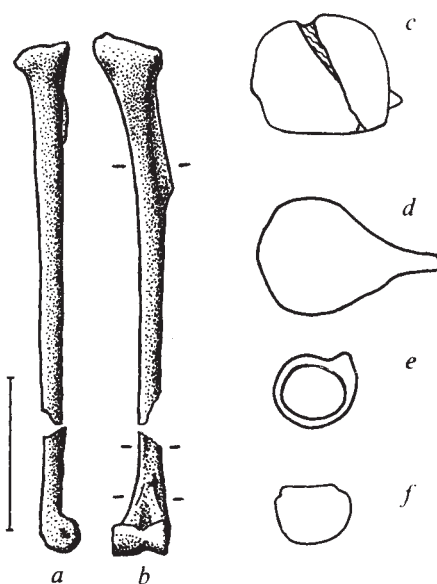


Fig. 1 The tibiotarsus (QM F12992) of *Nanantius eos*, gen. et sp. nov., in medial aspect (a) and anterior aspect (b). c, The head in proximal aspect; d, e, f, cross-sections of the shaft (with the extent of the internal cavity indicated only in e) at the three levels marked in b. Scale bar, 1 cm.

and species for this oldest-known enantiornithine. Hitherto, the only trace of birds from the Australian Mesozoic has been feathers found at Koonwarra, Victoria (Albian)⁴⁻⁶. These feathers are all small, about 2 cm long⁴⁻⁶ and hence are consistent, at least, with a bird about the size of that represented by the Hamilton tibiotarsus. Other previously reported bird material⁷ from the Australian Mesozoic seems to be dinosaurian.

The enantiornithine tibiotarsus is unique in the narrow intercondylar sulcus and unequal development of the distal condyles¹. The medial condyle is elongate while the lateral condyle is narrow, about half the length of the medial. No theropod dinosaur tibiotarsus (or astragalus) shows this: the condyles are more nearly equally developed, as in birds other than enantiornithines¹. The theropod tibia is distinguished by a short but prominent tibial crest. The Hamilton tibiotarsus has condyles and an intercondylar sulcus like those of the enantiornithines and lacks the tibial crest. Ornithischian dinosaurs have neither a tibiotarsus nor an ascending process of the astragalus, both found in the Hamilton tibiotarsus.

Those pterosaurs with a tibiotarsus have more evenly developed distal condyles which are more pulley-like in form⁸. The pterosaur astragalus has a distinct posterior groove⁹, and the epicondyles are offset in that a line drawn between them is inclined, not perpendicular, to the long axis of the element⁸. The tibia and fibula are fused in pterosaurs⁹. The unequal development of the distal condyles, epicondyles that are not offset, a separate tibia and fibula and the absence of a posterior groove indicate that this tibiotarsus is not derived from a pterosaur. These features indicate that it pertains to an enantiornithine, and hence to a bird.

Class Aves
Subclass Enantiornithes
Family?*

* No families have yet been formally proposed for the Enantiornithes; however, in view of the treatise on this group by C. A. Walker (manuscript in preparation) it would be inappropriate to propose families here.

Nanantius, gen. nov.

Type species: *Nanantius eos* sp. nov.

Type specimen: QM F12992, an isolated left tibiotarsus (Fig. 1).

Locality: North-east paddock of Warra Station, Queensland.

Stratigraphy: Toolebuc Formation.

Age: Albian.¹⁰

Diagnosis: A small enantiornithine, smaller than the smallest known from Salta and about the size of *Alexornis antecedens*, tibiotarsus more slender than any from Salta and similar in proportions to that of *Alexornis*, fibular crest relatively longer and stronger (Fig. 1d) than in the tibiotarsi from Salta, distinct ascending process of astragalus, marked depression anterior to fibular crest proximally, medial condyle more nearly cylindrical than in any other described enantiornithine.

Etymology: The generic name derives from the Greek *vavoc*, dwarf, and *ἑν-αντίος*, opposite, used here in reference to the subclass Enantiornithes. The specific name derives from the Greek *ἥως*, dawn, in reference to the earlier occurrence of this specimen than those of other enantiornithine taxa.

The tibiotarsus is preserved in two portions not sharing a contact, one 27 mm long, comprising the proximal two-thirds of the element, and the other, 9 mm long, the distal part. The entire tibiotarsus was probably no more than 40 mm in length. Comparison with the tibiotarsi from Salta clearly indicates that both pieces pertain to a single element. The subrectangular proximal articular face is lengthened transversely (Fig. 1c). The head is expanded, and overhangs anteriorly. The long, low fibular crest extends well down the shaft (Fig. 1b): proximally it merges with the head, leaving a marked concavity just below



Fig. 2 The tibiotarsus of *Nanantius eos* (QM F12992). a, Proximal portion in anterior view; b, distal condyles in medial view; c, in distal view. Not to scale.

the head and anterior to the crest (Fig. 2a). A much lower crest lies on the anterior side of the shaft, presumably a part of the anterior cnemial crest (Fig. 1a). The lateral cnemial crest is absent (as in the Salta tibiotarsi).

A low crest lies at the medial edge of the anterior face of the shaft, just distal to the break (Fig. 1e). The distal portion of the tibia is fused to the astragalus, leaving no discernible joint (Fig. 2b). The long axis of the astragal portion is nearly perpendicular to the long axis of the tibial shaft. The medial condyle is elongate and cylindrical in form, whilst the lateral condyle is narrower and with an articular surface formed like the basal part of a cone (Fig. 2c). The groove between the condyles is deeply V-shaped. The ascending process is broad and shallowly concave (Fig. 1f), with a low projection marking its apex. Breakage reveals that the shaft was hollow throughout, with thin walls (Fig. 1e).

The slender construction of the tibiotarsus may be related to the small size of this bone. However, the form of the medial condyle and distinctness of the astragal ascending process may not be. The latter character is probably plesiomorphic for birds, in view of the evidence for their descent from small theropod dinosaurs, and hence to be expected in an early member of the class.

The unique and recognizable structure of the enantiornithine tibiotarsus clearly apparent here indicates reference of this tibiotarsus to the Enantiornithes. This tibiotarsus differs from the enantiornithine tibiotarsi from Salta in having a narrower and relatively longer anterior cnemial crest and an elongate, rather than approximately square, proximal articular face. The Salta tibiotarsi have medial condyles that constrict slightly towards their medial terminations, while that of *Nanantius* does not. *Nanantius* also shows the fibular crest extending proximal

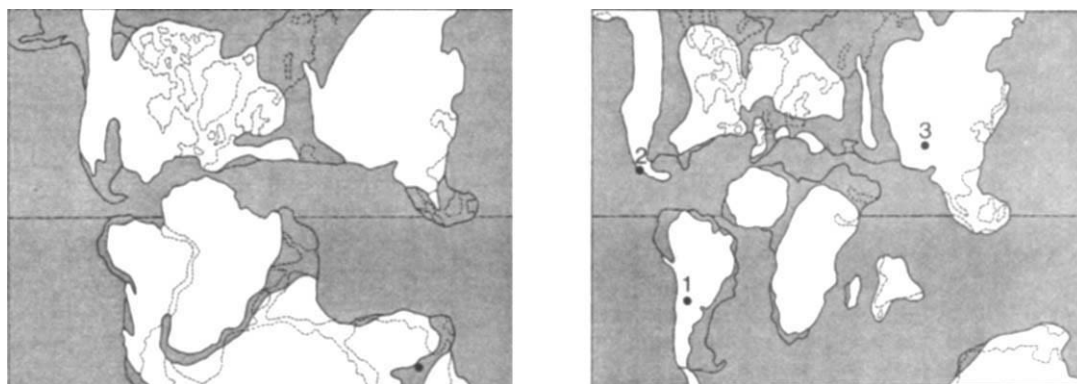


Fig. 3 Distribution of enantiornithines. Lower Cretaceous occurrences (left): Queensland, *Nanantius*. Upper Cretaceous occurrences (right): 1, Salta, *Enantiornis* and unnamed taxa; 2, Baja California, *Alexornis*; 3, Mongolia, *Gobipteryx*.

to the head and the astragalar ascending process, neither of which is apparent on the Salta material. The only preserved tibia of *Gobipteryx minuta* is very incomplete and from a hatching specimen¹¹, and hence cannot be satisfactorily compared with that of *Nanantius*. Only the proximal portion of the tibiotarsus of *Alexornis antecedens* is preserved¹², and this agrees in both size and proportions with that of *Nanantius*. However, the proximal part of the shaft in *Alexornis* is straight, whilst in *Nanantius* it curves medially (as is found in the Salta tibiotarsi). As in *Alexornis*, the fibular crest extends all the way up the shaft to the head, and the anterior and posterior faces of the crest are concave in section. The proximal internal articular surface is swollen, as in *Alexornis*, but shows a distinct posterior overhang lacking in that genus. Both cnemial crests are found in *Alexornis*, but only the anterior in *Nanantius*. Thus, *Nanantius eos* can be distinguished from the other enantiornithines known from comparable material, and this, together with its greater age, lends confidence to the conjecture that it may also be distinct from *Gobipteryx*.

Enantiornithines are known from the Maastrichtian of Argentina (*Enantiornis leali* and four other as yet undescribed taxa¹), the Campanian of Mexico (*Alexornis antecedens*¹²), the Campanian of Mongolia (*Gobipteryx minuta*¹¹) and the Albion of Australia. The reported occurrence of enantiornithine birds on four continents (Fig. 3), and ranging from the late early Cretaceous throughout the late Cretaceous, indicates that these birds were probably more widespread in the late Mesozoic than has hitherto been recognized. In addition, it emphasizes the widespread distribution of birds as a whole in the Lower Cretaceous, and provides another example of a bird in this period.

This discovery also suggests a link between the Australian Cretaceous vertebrate fauna and that of South America. Such a link has been suggested previously from the occurrence of allosaurid dinosaurs^{13,14} and similar ceratodontid lungfish¹⁵ on the two continents.

Much assistance has been provided in the study of this specimen by Drs N. Fraser, K. Campbell, P. H. V. Rich and R. Wild, and especially by Mr C. A. Walker, through whose courtesy I was able to examine undescribed material from Salta, Argentina. Mrs M. Norrie and Mr S. M. Van Dyck assisted in the collection and development of this material. Financial support of field work was provided by the Australian Research Grants Scheme.

7. Elzanowski, A. *Acta palaeont. pol.* **28**, 75-92 (1983).
8. Galton, P. M. *Palaeont. Z.* **54**, 331-342 (1980); Pls. 2M & 2O.
9. Padian, K. *Postilla* **189**, 1-44 (1983).
10. Senior, B. R., Mond, A. & Harrison, P. L. *Bull. Bur. Miner. Resour. Geol. Geophys. Aust.* **167**, 1-102 (1978).
11. Elzanowski, A. *Nature* **264**, 51-53 (1976); *Palaeont. pol.* **42**, 147-179 (1981).
12. Brodkorb, P. *Smithson. Contr. Paleobiol.* **27**, 67-83 (1976).
13. Molnar, R. E., Flannery, T. F. & Rich, T. H. *Alcheringa* **5**, 141-146 (1981); *J. Paleont.* **69**, 1513-1515 (1985).
14. Bonaparte, J. F. *Science* **205**, 1377-1379 (1979).
15. Kemp, A. & Molnar, R. E. *J. Paleont.* **55**, 211-217 (1981).

Oxytocin induces morphological plasticity in the adult hypothalamo-neurohypophyseal system

Dionysia T. Theodosia, Catherine Montagnese,
Floreal Rodriguez, Jean-Didier Vincent
& Dominique A. Poulain

INSERM, U176, Unite de Recherches de Neurobiologie des Comportements, Rue Camille St Saëns, Bordeaux, F33077 Cedex, France

The hypothalamo-neurohypophyseal system offers a unique example in the adult mammalian central nervous system (CNS) of a functional and structural plasticity related to a physiological state. During lactation, oxytocin neurones evolve a synchronized electrical activation which permits pulsatile hormone release at milk ejection¹. At the same time, in the supraoptic (SON) and paraventricular nuclei, glial coverage of neurones diminishes, so that large portions of their surface membrane become directly juxtaposed; synaptic remodelling also associates pairs of neurones through the formation of common presynaptic terminals²⁻⁵. These structural changes, reversible after weaning⁴, affect exclusively oxytocinergic neurones⁵ and could facilitate their synchronized electrical activity. As several observations suggest that oxytocin itself is released centrally⁶⁻⁸, we have examined the effect of prolonged intracerebroventricular infusions of oxytocin on the structure of the SON of non-lactating animals. We report here that the peptide indeed engenders the structural reorganization characteristic of the oxytocin system when it is physiologically activated. Similar infusion of vasopressin has no effect. Our observations thus demonstrate that a central neuropeptide can induce anatomical changes in the adult CNS, and suggest that oxytocin can regulate its own release by contributing to the dramatic restructuring of the nuclei containing the neurones responsible for its secretion.

Our experiments were performed on primiparous Wistar rats from our colony 1 month after weaning their first litter. The

Received 27 March; accepted 16 May 1986.

1. Walker, C. A. *Nature* **292**, 51-53 (1981).
2. Martin, L. D. in *Perspectives in Ornithology*, 291-338 (Cambridge University Press, 1983).
3. Molnar, R. E. & Thulborn, R. A. *Nature* **288**, 361-363 (1980).
4. Talent, J. A., Duncan, P. M. & Handby, P. L. *The Emu* **66**, 81-86 (1966).
5. Waldman, M. *Condor* **72**, 377 (1970).
6. Rich, P. V. *Proc. 16th int. Congr. 53-65* (Australian Academy of Science, 1976).

Fig. 1 Electron micrographs of oxytocinergic neurones in the supraoptic nucleus (SON) of rats given an intracerebroventricular infusion of oxytocin for 8 days. In *a*, the surface membranes of two soma profiles are seen to be in extensive apposition, that is, they are not separated by any neuropile or glial element (small arrows). Only a portion of the 15- μ m-long apposition is shown in this micrograph. The oxytocinergic nature of the somata is revealed by colloidal gold particles (10 nm) over their secretory granules (see also insets), after immunostaining with a rabbit anti-oxytocin serum and immunoglobulin-coupled colloidal gold. In *b*, a presynaptic terminal (asterisk) synapses (between arrowheads) onto two adjacent soma profiles that are also juxtaposed (ap, attachment plaque bridging the apposed surface membranes). Colloidal gold particles (5 nm) over secretory granules in their cytoplasm identify the somata as oxytocinergic (see also inset). Scale bars, 0.5 μ m; insets, 0.1 μ m.

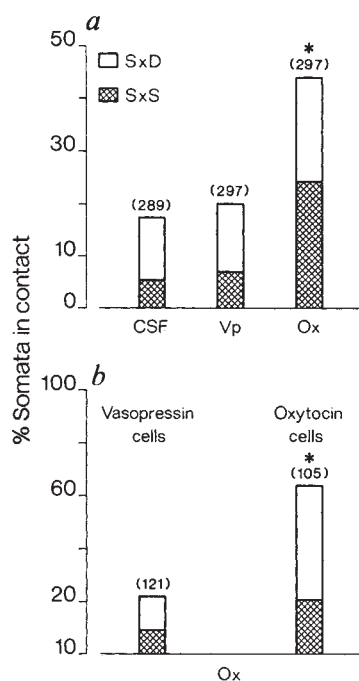
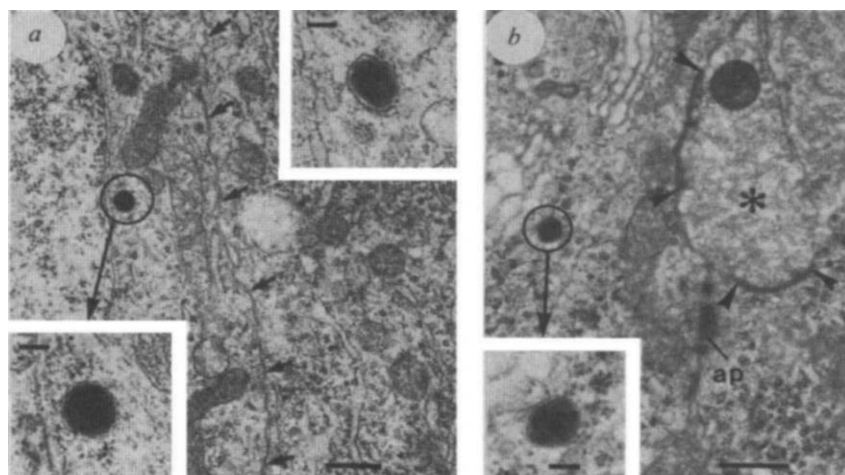


Fig. 2 *a*, Percentage of neurosecretory soma profiles showing neuronal appositions in the SON of rats after chronic (8 days) intracerebroventricular infusion of artificial cerebrospinal fluid (CSF), vasopressin (Vp) or oxytocin (Ox). This analysis was carried out without immunocytochemical identification of the neurones. Following infusion of oxytocin, nearly half of all the observed neurosecretory soma profiles were directly apposed to adjacent somata (S×S) or dendrites (S×D). *b*, Percentage of oxytocinergic and putative vasopressinergic soma profiles, identified immunocytochemically using anti-oxytocin serum, involved in S×S and S×D contacts in the SON of rats given intracerebroventricular infusions of oxytocin. In both histograms, numbers in parentheses represent the total number of soma profiles examined in each group. * $P < 0.001$, χ^2 analysis on raw data.

Methods. Two different quantitative analyses were carried out by two separate investigators, unaware of the experimental conditions. In the first (*a*), the analysis was performed on electron micrographs of neurosecretory soma profiles selected by a random sampling procedure (for further details see ref. 2); for each group, each composed of 5 animals, 128 micrographs were examined. In the second analysis (*b*), counting was performed directly at the electron microscope. In order to examine in detail larger areas of tissue surrounding each neurosecretory soma profile, all tissue falling within the whole surface of randomly chosen grid squares covering the whole frontal span of the SON was examined⁵. Four sections, each cut from different anterior-posterior levels of the SON of each of four animals that had been infused with oxytocin, were examined; the sections had been immunostained with a rabbit anti-oxytocin serum (preabsorbed with vasopressin adsorbed to CNBr-activated Sephadex beads) and immunoglobulin-coupled colloidal gold (a post-embedding immunocytochemical procedure described fully elsewhere⁵). The same parameters were noted in each analysis: (1) total number of neurosecretory soma profiles; in the case of the immunostained SON, the peptide content of the somata was also noted (see Fig. 1); (2) number of soma profiles in juxtaposition to an adjacent soma or dendritic profile; (3) number of dendritic contacts per soma profile; and (4) number of 'double' synapses making contact with two neurosecretory soma or dendritic profiles simultaneously.

animals were deeply anaesthetized with sodium pentobarbital (50 mg per kg), and a cannula, connected to an Alzed miniosmotic pump (Scientific Marketing), delivering fluid at a rate of 0.5 μ l h⁻¹, was inserted stereotactically into the third ventricle. The pump was then left under the skin of the back at the level of the shoulders. The pumps contained synthetic oxytocin or vasopressin (Peninsula, 2 μ g ml⁻¹ dissolved in artificial cerebrospinal fluid (CSF)), or artificial CSF alone. Eight days later, the animals were anaesthetized and perfused intracardially with heparinized saline followed by fixative (4% paraformaldehyde and 2.5% glutaraldehyde in 0.1 M sodium phosphate buffer). The brains were removed and 100- μ m coronal slices were cut on a vibratome. The position of the cannula was verified on these sections. Each pump was emptied and the remaining fluid volume measured. The concentration of oxytocin and vasopressin remaining at the end of each infusion was measured by radioimmunoassay⁹. The biological activity of the remaining oxytocin was also tested by its potency to induce an increase in intramammary pressure in mammary glands of lactating rats. Once it was ascertained that the cannula had been correctly

placed, and that the pumps had functioned properly, blocks containing the SON were dissected from the vibratome slices and processed further for electron microscopy (see ref. 5).

In the SON of animals infused with artificial CSF or vasopressin, as in nuclei from unstimulated animals (hydrated male^{2,10,11}, virgin female^{2,4} or post-weaned rats⁴), most neurosecretory soma profiles were separated from other neuronal profiles by glial or neuropile elements. On the other hand, in the SON of rats infused with oxytocin, glial coverage diminished markedly and close to 45% of all neurosecretory soma profiles were directly juxtaposed to adjacent soma or dendritic profiles (Figs. 1*a*, 2*a*). Moreover, the number of neuronal contacts more than doubled in this group, from 0.2 to 0.7 soma or dendritic contacts per soma profile. Concurrently, many of the same neurones (13%) were bridged by the same presynaptic terminal ('double' synapses; Fig. 1*b*), a type of synapse rarely seen in the nuclei of untreated animals (<5%^{2-4,11}) or of animals infused with vasopressin or CSF (6%). In order to identify the neurones, a separate analysis was carried out on sections of the SON that had been immunostained for

oxytocin (Fig. 1a,b), using a post-embedding immunocytochemical procedure described in detail elsewhere⁵. This revealed that most of the soma profiles showing surface membrane appositions and sharing the same synapse after oxytocin infusion were oxytocinergic (Figs 1, 2b).

Our experiments demonstrate that a sustained elevated level of oxytocin in the third ventricle leads to morphological changes in the SON which are identical in nature and extent to those that occur under physiological^{2,4,5,11} or experimental^{10,11} conditions when the oxytocinergic system is activated. Our observations strongly suggest, therefore, that oxytocin does indeed take part in the restructuring of the hypothalamo-neurohypophysial system. Moreover, since infusion of the structurally similar neuropeptide vasopressin had no effect, it seems that oxytocin is specifically responsible for the anatomical changes.

Under normal conditions, oxytocin is present in the CSF¹² but its origin remains unclear. It may be released from oxytocin-containing neurones located in many areas of the CNS¹³. In particular, recent observations have described a local, calcium-dependent release of this peptide from the magnocellular nuclei^{6,7}, and oxytocinergic terminals that are in contact with oxytocin cell bodies and dendrites have been noted in the SON⁸. Unlike the blood-brain interface, there is no barrier to the diffusion of oxytocin between the third ventricle and neighbouring cerebral structures¹². Our intracerebroventricular infusions of oxytocin would thus mimic an enhanced central release of oxytocin. The question then arises as to how a centrally circulating neuropeptide can induce anatomical changes in a particular

region of the CNS. Central oxytocin has been shown to increase the electrical activity of oxytocin neurones¹⁴ and the anatomical changes could be a consequence of such an activation. This possibility is supported by the observation that restructuring of the nuclei occurs under other situations, such as parturition^{3,4}, and chronic dehydration^{10,11}, which also enhance the overall activity of oxytocinergic neurones¹. It is also possible, however, that oxytocin itself is the signal, acting on the surface membranes of oxytocin neurones, their synapses, or surrounding glia by an unknown mechanism.

The anti-oxytocin serum was generously provided by Dr A. Burlet. We also thank Mme A. Bonhomme for technical assistance and Mr S. Senon for photographic work.

Received 14 April; accepted 12 June 1986.

1. Poulain, D. A. & Wakerley, J. B. *Neuroscience* **7**, 773-808 (1982).
2. Theodosis, D. T., Poulain, D. A. & Vincent, J. D. *Neuroscience* **6**, 919-929 (1981).
3. Hatton, G. I. & Tweedle, C. D. *Brain Res. Bull.* **8**, 197-204 (1982).
4. Theodosis, D. T. & Poulain, D. A. *Neuroscience* **11**, 183-193 (1984).
5. Theodosis, D. T., Chapman, D. B., Montagnese, C., Poulain, D. A. & Morris, J. F. *Neuroscience* **17**, 661-678 (1986).
6. Moos, F. et al. *J. Endocr.* **102**, 63-72 (1984).
7. Mason, W. T., Hatton, G. I., Ho, Y. W., Chapman, C. & Robinson, I. C. A. F. *Neuroendocrinology* **42**, 311-322 (1986).
8. Theodosis, D. T. *Nature* **313**, 682-684 (1985).
9. Cazalis, M., Dayanithi, G. & Nordman, J. J. *J. Physiol., Lond.* **369**, 45-60 (1985).
10. Tweedle, C. D. & Hatton, G. I. *Brain Res.* **309**, 373-376 (1984).
11. Chapman, D. B., Theodosis, D. T., Montagnese, C., Poulain, D. A. & Morris, J. F. *Neuroscience* **17**, 679-686 (1986).
12. Robinson, I. C. A. F. *Prog. Brain Res.* **60**, 129-145 (1983).
13. Sofroniew, M. V. in *Handbook Chemical Neuroanatomy* Vol. 4, Pt 1 (eds Bjorklund, A. & Hokfelt, T.) 93-165 (Elsevier, Amsterdam, 1985).
14. Freund-Mercier, M. J. & Richard, P. J. *J. Physiol., Lond.* **352**, 447-466 (1984).

Serine esterase in cytolytic T lymphocytes

Mark S. Pasternack*†‡, C. Reynold Verret*, Margaret A. Liu*† & Herman N. Eisen*

* Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

† Children's and ‡ Medical Services, Massachusetts General Hospital, Boston, Massachusetts 02114, USA and Departments of Medicine and Pediatrics, Harvard Medical School, Boston, Massachusetts 02115, USA

The mechanisms that enable cytotoxic T lymphocytes (T_c cells) to destroy target cells are only vaguely understood. However, recent studies have identified in T_c cells¹ and natural killer cells² cytoplasmic granules that contain perforin, a cytolytic protein that resembles the ninth component of complement (C9)³. Antigen-specific lysis of target cells, traditionally ascribed solely to T_c cells, has now also been demonstrated in some T-helper cell (T_h cell) lines, referred to here as T helper-killer or $T_{h/c}$ cells⁴⁻⁶. We recently found a novel serine esterase that is present at greatly elevated levels in cloned murine T_c cell lines and one $T_{h/c}$ cell line, but not in two non-cytolytic T_h cell lines⁷. These findings suggest that the serine esterase is involved in cytolytic activity and that a variety of effector cells share a common cytolytic mechanism. To explore the role of the serine esterase in this process, we have been studying additional properties of the enzyme in murine T cells. We show here that (1) it is a membrane-associated, disulphide-linked dimer, (2) it has trypsin-like properties but is not a general protease, (3) in density gradient centrifugation it sediments with perforin, (4) it is secreted by T_c cells during their cytolytic attack on target cells, and (5) antiserum to T_c -cell serine esterase reacts with the enzyme in $T_{h/c}$ cells.

The serine esterase was isolated from T_c clone G4 cells as described in Fig. 1 legend, and the isolated protein was analysed by SDS-polyacrylamide gel electrophoresis (PAGE) under

reducing and nonreducing conditions. Under both conditions the same complex was observed, but their apparent relative molecular mass (M_r) differed by a factor of about two (31,000-34,000 reduced and 55,000-60,000 nonreduced). Exactly the same difference in M_r was seen under these conditions when the enzyme was visualized by subjecting unfractionated cell lysates to affinity labelling with ³H-diisopropylphosphofluoridate (³H-DFP)⁷ and autoradiography (data not shown). It appears, therefore, that the enzyme is a disulphide-linked dimer. Whether its evident microheterogeneity reflects variable glycosylation or the presence of a few distinct but related ³H-DFP-reactive serine esterases is not clear.

To determine whether all of the disulphide-linked subunits have DFP-reactive sites, the enzyme was examined under reducing conditions by two-dimensional PAGE (using non-equilibrium pH gradient electrophoresis (NEPHGE) in the first dimension)^{8,9}, visualizing the protein in some gels by silver stain¹⁰ (Fig. 1c) and in others by affinity labelling with ³H-DFP (Fig. 1d). In both cases, the appearance of the reduced protein was the same: a predominant component of M_r ~31,000 located close to the basic edge of the pH gradient with some heterogeneity evident in the form of a trail of slightly less basic protein of slightly higher M_r (32,000-34,000) (Fig. 1c, d). A single component, like that of Fig. 1c and d, was also seen following very brief NEPHGE electrophoresis (800 Vh, data not shown), indicating that another, more basic subunit had not been overlooked. Because the appearance of the reduced and nonreduced enzyme was the same when visualized by silver stain and ³H-DFP, both subunits in each dimer must be labelled by DFP, and have the same or very similar net positive charge; they are thus very similar and possibly identical.

To facilitate comparison of serine esterase in diverse clones of cytotoxic cells, rabbit antisera were raised against the purified enzyme (Fig. 1). The antisera immunoprecipitated only one ³H-labelled component from ³H-DFP-treated lysates of G4 cells (data not shown) and they also reacted with the unlabelled lysates of cytolytic T cells by Western blotting (Fig. 2). In both assays the reactive material evidently corresponded to the serine esterase, as it was indistinguishable in M_r and appearance from the ³H-DFP affinity-labelled and silver-stained enzyme seen in

Fig. 1 *a*, Purification of serine esterase from a clone of T_c cells, monitored by enzyme activity (absorbance at 412 nm) and by SDS-PAGE (insert), using silver impregnation to visualize protein. Insert: Lane A, the eluate from a lentil-lectin column (see Methods below) loaded on the lysine-Sepharose column; B–N, fractions from the lysine-Sepharose column: B, fraction 4; C, fraction 7; D, fraction 10; E, fraction 13; F, fraction 16; G, fraction 19; H, fraction 22; I, fraction 25; J, fraction 28; K, fraction 31; L, fraction 34; M, fraction 37; N, fraction 40; O, M_r markers ($\times 10^{-3}$). *b–d*, Some physical properties of the T-cell serine esterase. *b*, SDS-PAGE of purified esterase (pooled fractions 25–40, *a*) stained by silver impregnation. M, M_r markers (top to bottom $\times 10^{-3}$): 90, 67, 45, 31 and 17. NR, nonreduced; R, reduced. *c*, Two-dimensional electrophoresis under reducing conditions of 5 μ g of purified serine esterase. The first dimension was non-equilibrium pH gradient electrophoresis (NEPHGE) and the second dimension was SDS-PAGE; protein was visualized by silver impregnation. *d*, Two-dimensional electrophoresis as in *c* of an Nondet P-40 (NP40) extract of T_c cells (clone G4) labelled with ³H-DFP. The DFP-reactive protein was visualized by autoradiography of the slab gel.

Methods. *a*, $1\text{--}2 \times 10^9$ cells of T_c clone G4 (BALB.B anti-D^d, ref. 19) were suspended in lysis buffer (135 mM NaCl, 10 mM HEPES pH 7.4, 3.5 mM MgCl₂) at $\sim 10^8$ cells ml⁻¹ and disrupted by nitrogen cavitation (300 p.s.i. 15 min, 4 °C). Nuclei and whole cells were pelleted by centrifugation at 300g, 5 min. A wash of the pellet with lysis buffer was added to the post-nuclear supernatant, which was centrifuged at 100,000g for 1 h at 4 °C. The resulting supernatant (which had <1% of the initial esterase activity) was discarded and the pelleted membranes were suspended in 150 mM NaCl, 10 mM Na phosphate pH 7.4, 0.5% NP40 and 0.02% NaN₃ (PBS/NP40) and loaded onto a lentil-lectin Sepharose column (Pharmacia) equilibrated in PBS/NP40 and washed with additional buffer. Approximately 60% of the activity was not retained and treatment of the pass-through with ³H-DFP revealed a single sharply defined band of M_r 31,000 on SDS-PAGE (not shown). The remaining 40% of esterase activity, retained on the lentil-lectin column, was recovered by elution with 0.5 M α -methyl mannoside (α -MM) in PBS/NP40. After dialysing this eluate against 15 mM Tris-HCl pH 8.1, 0.02% NaN₃, it was loaded (*a*, insert) onto a 5-ml lysyl-Sepharose column (Pharmacia) previously equilibrated with 15 mM Tris-HCl pH 8.1, 0.5% NP40, 0.02% NaN₃ (TNA) and washed with this buffer. The column was monitored by assaying 1 μ l of each fraction (~ 1.0 ml) for esterase activity, using benzoyloxycarbonyl lysyl thioester (BLT) (Calbiochem) as described previously⁷. No esterase activity was detected in the pass-through. A NaCl gradient (0–500 mM in TNA) eluted the activity as a single peak, as shown, with the peak height at about 250 mM NaCl. For some preparations the column was equilibrated and washed with 1% octylglucoside in 15 mM Tris-HCl pH 8.1, before applying the NaCl gradient. Aliquots (20 μ l) of fractions from the lysyl-Sepharose column were analysed under reducing conditions by electrophoresis in a 10% acrylamide gel in the presence of 0.1% SDS²⁰. After fixation of the slab gel in 50% methanol–10% acetic acid, the proteins were visualized by silver impregnation¹⁰. Fractions 25–40 were pooled. The serine esterase activity in the pool usually represented $\sim 20\%$ of the total initial activity and $\sim 70\%$ of the activity added to the lysine-Sepharose column. *b*, Purified esterase (~ 2.5 μ g) was heated to 100 °C for 3 min in sample buffer not containing (NR) or containing (R) 5% 2-mercaptoethanol and analysed by electrophoresis in a 10% polyacrylamide gel in the presence of 0.1% SDS²⁰. Silver impregnation was performed as described elsewhere¹⁰. *c*, Purified esterase (~ 5 μ g) was dissolved in 25 μ l of isoelectric focusing sample buffer containing 5% 2-mercaptoethanol, applied to a 4% polyacrylamide gel (2.5 mm $\times$ 12.5 mm) and electrophoresed under non-equilibrium conditions^{8,9} for 1,400 Vh. The gel was incubated with SDS sample buffer containing 5% 2-mercaptoethanol and layered on a 10% polyacrylamide slab gel and run in the presence of 0.1% SDS²⁰. Silver impregnation was performed as described elsewhere¹⁰. *d*, An extract of 3×10^6 G4 cells in PBS containing 0.5% NP40 was treated with 10^{-5} M ³H-DFP for 30 min at 37 °C and prepared for electrophoresis as described elsewhere⁷. NEPHGE (1,300 Vh) and SDS-PAGE were performed as for *b*. The gel was impregnated with Autofluor (National Diagnostics, Somerville, New Jersey), dried and exposed for 8 days at -70 °C using XAR-5 film (Eastman Kodak).

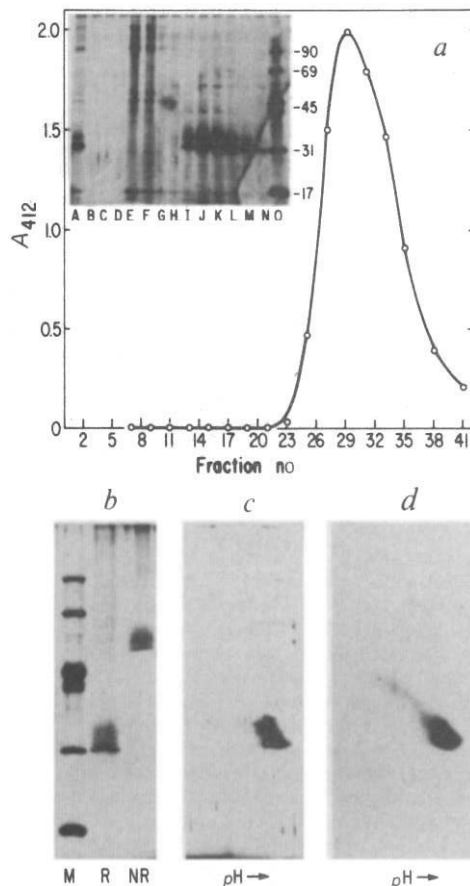
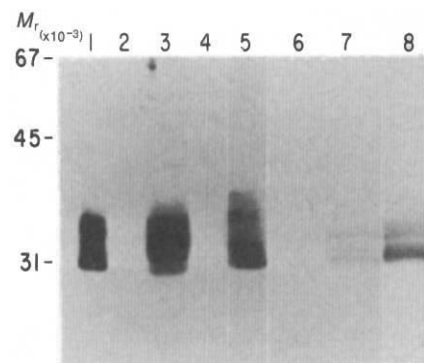


Fig. 2 Western blot analysis of cloned T_c, T_h and T_{h/c} cell lines. Lane 1, purified serine esterase (Fig. 1); lane 2, T_h cells (clone TDH-1); lane 3, T_{h/c} cells (clone 5.5); lane 4, T_h cells (clone D10); lane 5, T_c cells (clone 2C); lane 6, normal thymocytes; lane 7, Con-A-activated thymocytes; lane 8, T_c cells (clone G4). M_r markers at the left are bovine serum albumin (67,000), ovalbumin (45,000), carbonic anhydrase (31,000).

Methods. Antiserum to the serine esterase was raised by injecting rabbits with ~ 100 μ g of purified enzyme (Fig. 1) in complete Freund's adjuvant and boosting them twice at 1–2-month intervals with the same amount of the enzyme in incomplete adjuvant. Cell extracts were prepared by washing cells several times in PBS and incubating them in 0.5% PBS/NP40 at 0 °C for 30 min with frequent vortexing. SDS was added to 0.1% and incubation was continued for another 15 min. After centrifuging the lysates at 8,000g for 10 min, supernatants ($1\text{--}4 \times 10^6$ cell-equivalents) were subjected to electrophoresis in SDS-10% polyacrylamide gels. Proteins were transferred electrophoretically from gels to nitrocellulose or Zeta-bind filters (A.M.F. Cuno) and, to reduce nonspecific binding in subsequent steps, the filters were incubated (1 h, room temperature) in 5% low-fat milk (Carnation) in PBS or in a 1:5 dilution of fetal calf serum in TBS (155 mM NaCl, 10 mM Tris-Cl pH 7.5) (TBS-FCS). The filters were then incubated for at least 3 h in anti-serine esterase antiserum diluted 250-fold in PBS-low fat milk or TBS-FCS, washed copiously with 0.05% Tween 80 in TBS, and finally treated sequentially with biotinylated anti-rabbit globulin and an avidin complex with biotinylated horseradish peroxidase, as described previously²¹. Colour was developed with 4-chloro-1-naphthol and hydrogen peroxide as described elsewhere²¹. Clone TDH-1 is an L3T4⁺, 2,4,6 trinitrophenyl + H-2^d-specific clone of BALB/c origin; clone 5.5 is an L3T4⁺, ovalbumin + I-E^d-specific clone of BALB/c origin⁴; D10 is a conalbumin + I^k-specific clone of AKR origin⁵; clone 2C is BALB.B anti-L^d (ref. 22); activated thymocytes were generated as described previously^{7,11}; G4 is described in Fig. 1 legend.



other gels (Fig. 1 and ref. 7). The Western blots (Fig. 2) showed the enzyme to be present at a high level in three cytolytic cell lines (two T_c and one T_{h/c} cell line: 2C, G4 and 5.5, respectively) but to be undetectable in two non-cytolytic (T_h) cell lines (TDH-1 and D10). The antisera also detected the esterase in Western blots of murine thymocytes, providing these cells had been previously stimulated for 4 days in culture by concanavalin A (Con-A, Fig. 2, lanes 6, 7); these conditions induce the

expression of both cytotoxic¹¹ and serine esterase activity⁷. The antiserum to the enzyme from the T_c G4 cells reacted with enzyme in all of the other T_c clones tested and also in a T_{h/c}-cell line. The slight variations in M_r values in different T-cell lines, and in the Con-A-activated thymocytes, suggest that there may be several related T-cell serine esterases or that a single enzyme is modified differently (for example, by glycosylation) in different T-cell lines.

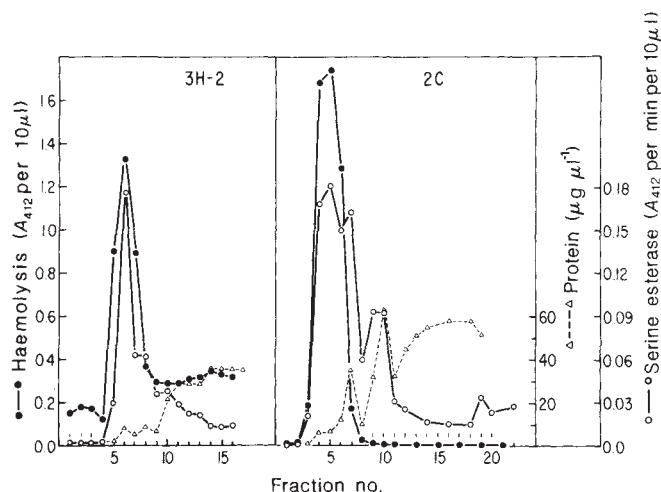


Fig. 3 Percoll-gradient fractionation of two cloned T_c -cell lines. *a*, Clone 3H2; *b*, clone 2C.

Methods. 5×10^8 3H2 cells (BALB.B anti-D^d, Lys-2⁺, see ref. 23) and 2.5×10^8 2C (ref. 22) cells were washed twice in Hank's balanced salt solution and once in 'relaxation' buffer (100 mM KCl, 3 mM NaCl, 3.5 mM MgCl₂, 1.25 mM EGTA, 1 mM ATP, 10 mM PIPES pH 6.8)²⁴. The cells were resuspended in relaxation buffer and lysed by nitrogen cavitation at 350 p.s.i. for 30 min. Nuclei and debris were removed by centrifugation at $\sim 1,000$ g for 5 min, and the supernatant was applied to a discontinuous gradient of Percoll in relaxation buffer (4 ml 90% Percoll, 2 ml 60%, 3.5 ml 39%) and subjected to centrifugation at 20,000 r.p.m. in a Beckman SW27 rotor for 30 min. Fractions containing ~ 0.8 ml were collected from the bottom of the tube with a peristaltic pump. Assay for haemolytic activity: The assay is a modification of Masson and Tschopp's procedure¹². Sheep erythrocytes were washed in HEPES buffered saline (HBS: 155 mM NaCl, 10 mM HEPES pH 7.3) and suspended at $\sim 3 \times 10^8$ cells ml⁻¹ in HBS containing 5 mM CaCl₂. Diluted or undiluted fractions (10 µl) were added to 100 µl of a sheep erythrocyte suspension. After 15 min at 37 °C, 1 ml of PBS containing 1 mM EGTA was added and the mixture was centrifuged at 1,500g for 5 min. Haemoglobin released in the supernatant was determined from the absorbance at 412 nm. Esterase activity was determined from plots of A_{412} against time. Protein was quantitated with the BCA protein assay reagent (Pierce) using bovine serum albumin standard. Percoll was removed after colour development by centrifugation at 8,000g for 5 min. Background due to Percoll and relaxation buffer was negligible.

All nine T_c cell lines tested showed markedly elevated enzyme levels: for example, 50–200 U per 10^6 cells⁷, corresponding at the upper limit to 1 µg trypsin-equivalents per 10^6 cells⁷. The enzyme level was also elevated in two $T_{h/c}$ -cell lines (clones 5.5 and 18) but not in four non-cytolytic T_h -cell lines (O₃, D₅ (ref. 7), D10 and TDH-1). Although the presence of elevated serine esterase thus seems to correlate with cytolytic activity, many more murine T-cell lines, especially T_h and $T_{h/c}$ lines, will have to be examined to determine how consistently the correlation holds up.

The cytolytic protein perforin has been identified in cytoplasmic granules of human and mouse natural killer cells and mouse T_c -cell lines^{1,2}. By virtue of their high density, the granules can be separated from other cell lysate components and their isolation can be monitored by their lysis of sheep red blood cells. As Fig. 3 shows, the haemolytic and serine esterase activities sedimented together in T_c -cell lysates in dense fractions near the bottom of the Percoll gradient. This apparent association of the serine esterase with the cytolytic granules is supported by a recent analysis of M_r values of the reduced and nonreduced proteins isolated from the dense granules of a murine T_c -cell line¹² (compare our Fig. 1b with Fig. 2c in ref. 12) and by our unpublished observations that the ³H-DFP affinity-labelled pro-

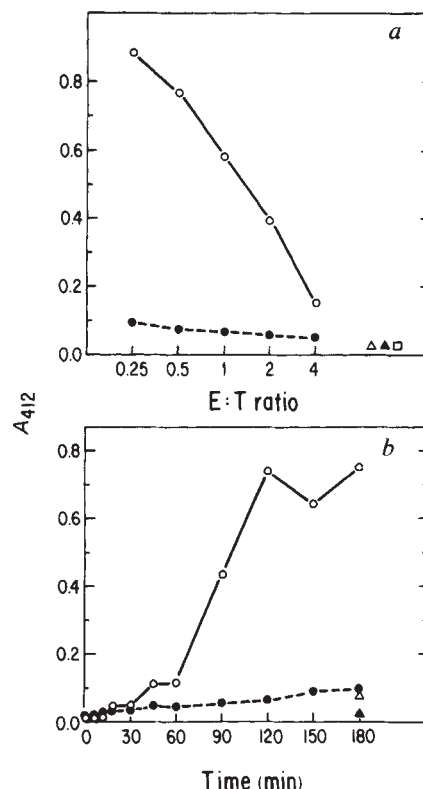


Fig. 4 Serine esterase secretion is triggered by cytolytic attack of T_c cells on specific target cells. *a*, Influence of the ratio of effector (T_c) cells to target cells (E:T ratio). *b*, Time course of serine esterase secretion. ○, Cells of T_c clone G4 (anti-D^d) incubated with P815 (H-2^d) target cells; ●, G4 cells incubated EL-4 (H-2^b) cells; △, G4 cells incubated alone; ▲, P815 cells incubated alone; □, EL4 incubated alone.

Methods. G4 cells¹⁹ were collected, washed, resuspended at 5×10^5 cells ml⁻¹ in complete medium containing 20 U ml⁻¹ recombinant interleukin-2 and distributed into flat-bottomed microtitre wells (Costar 3596) at 10^5 cells per well. G4 cells readily attached to the plastic surface. After overnight incubation, the wells were gently aspirated and filled with cold KBI medium (complete medium supplemented with bovine serum albumin (Boehringer Mannheim) at 1 mg ml⁻¹ in place of fetal calf serum). After several minutes on ice, the wells were aspirated again and to them were added (in triplicate) 100 µl medium alone or 100 µl medium containing varying numbers of P815 or EL-4 cells which had been washed two or three times in KBI medium and resuspended in KBI at the designated concentration. The plates were incubated at 37 °C for 3 h in *a* or for the time noted in *b*. Samples (50 µl) were removed from replicate wells and pooled. Esterase activity in 10 µl of the pooled supernatants was measured as described elsewhere⁷. In *a*, E:T ratios were based on the nominal number of T_c cells seeded before the assay. Recovery experiments showed that no cell proliferation had occurred during overnight culture (data not shown). In *b*, the E:T ratio was held constant at 0.5 and plates were centrifuged briefly (500 r.p.m., 45 s) at 4 °C before initiation of the 37 °C incubation. The use of adherent T_c cells facilitated the detection and measurement of secreted esterase.

tein is present in the dense fractions that possess haemolytic activity (Fig. 3). Taken together, these findings strongly suggest that in murine T_c cells the serine esterase is present in the dense cytotoxic granules that contain perforin.

The serine esterase was previously shown⁷ to be blocked by DFP and phenylmethylsulphonyl fluoride, typical inhibitors of serine proteases. It is also blocked by aprotinin (data not shown), another typical inhibitor of trypsin-like serine proteases. However, the serine esterase did not cleave ¹²⁵I-casein (data not shown), suggesting that it may have restricted proteolytic specificity, as is characteristic of regulatory proteases¹³ (for

example, in some of the sequentially acting components of the complement 'cascade'). Whether the substrate for the serine esterase is perforin, or another component of the perforin-containing granules, or perhaps a surface protein of T_c or target cells remains to be determined.

T_c cells normally lyse only cells whose surface antigens they recognize, not neighbouring cells that lack these antigens, suggesting that target cell lysis by T_c cells is not due to secretion of a diffusible cytotoxic molecule. However, when clone G4 cells were incubated with target cells, serine esterase activity appeared in the extracellular medium (Fig. 4). The enzyme released during lysis of the target cells remained in the supernatant after low-speed centrifugation (300 g, 5 min), sufficient to sediment intact cells, but was almost entirely (95%) sedimented following centrifugation at 100,000 g for 1 h. Under the latter conditions the purified enzyme (Fig. 1) is not sedimented, and thus the released enzyme was probably still largely associated with granule membranes or with another granule component, perhaps a proteoglycan^{14,15}. Enzyme release clearly accompanied the cytolytic attack on H-2^d antigen-bearing target cells (P815), since it was not observed (Fig. 4) in various controls, for example, G4 cells alone or G4 cells incubated with cells possessing the wrong H-2 haplotype (EL-4 cells, H-2^b).

The amount of enzyme secreted varied with the ratio of effector (T_c) cells to target cells (E:T ratio). With an excess of target cells (low E:T ratio) the amount secreted was high, probably because under these conditions almost every T_c cell was likely to be interacting with a target cell; with a scarcity of target cells (high E:T ratios) the amount secreted was low, probably because under these conditions many T_c cells were not engaged with target cells (Fig. 4a).

The time course of secretion is shown in Fig. 4b. About two-thirds of the amount present initially in the T_c cells was released. Whether enzyme biosynthesis keeps pace with secretion, or whether secretion results in sufficient depletion to reduce cytolytic activity ('killer cell exhaustion') are among the many issues that remain to be investigated. Target cell-triggered secretion of γ -interferon¹⁶ and interleukin-2 (ref. 17) from T_c cells has been observed previously. Comparison of the slow rate of secretion of γ -interferon¹⁸ with the far more rapid rate for the serine esterase suggests that the encounter of T_c with target cells induces the T_c cells to synthesize and then secrete γ -interferon, whereas the same stimulus triggers the T_c cells to release their preformed serine esterase, contained in the cytolytic granules. Thus, in addition to their cytolytic effects the release of this enzyme and perhaps additional granule components by T_c and $T_{h/c}$ cells might contribute to the slowly evolving local tissue injury known as delayed-type hypersensitivity.

We thank the following for providing cloned T-cell lines that have not yet been described in publications: David M. Kranz for $T_{h/c}$ clone 18, Edward B. Reilly for T_c clone 3H2, and George Sigal for T_h clone TDH-1. We also thank Charles Janeway and John Tite for $T_{h/c}$ clone 5.5 and T_h clone D10, Robert D. Rosenberg and James Marcum for valuable discussions and purified casein, and Ann Hicks for preparation of this manuscript. This work was supported by research grants (CA-15472 and CA-28900) and a Cancer Center Core Grant (CA-14051) from the National Cancer Institute, NIH. C.R.V. is the recipient of a Massachusetts Institute of Technology Minority Fellowship Award and M.A.L. is the recipient of a Physician-Scientist Award from the NIH.

7. Pasternack, M. S. & Eisen, H. N. *Nature* **314**, 743-745 (1985).
8. O'Farrell, P. Z., Goodman, H. M. & O'Farrell, P. H. *Cell* **12**, 1133-1142 (1977).
9. Jones, P. P. in *Selected Methods in Cellular Immunology* (eds Mishell, B. B. & Shiigi, S. M.) 398-440 (Freeman, San Francisco, 1980).
10. Morrissey, J. H. *Analyt. Biochem.* **117**, 307-310 (1981).
11. Irlé, C., Piquet, P. F. & Vassalli, P. *J. exp. Med.* **148**, 32-45 (1978).
12. Masson, D. & Tschopp, J. *J. biol. Chem.* **260**, 9069-9072 (1985).
13. Neurath, H. *Fedn Proc.* **44**, 2907-2913 (1985).
14. Schmidt, R. E. *et al. Nature* **318**, 289-291 (1985).
15. MacDermott, R. P. *et al. J. exp. Med.* **162**, 1771-1787 (1985).
16. Klein, J. R., Raulat, D. H., Pasternack, M. S. & Bevan, M. J. *J. exp. Med.* **155**, 1198-1203 (1982).
17. Andrus, L., Granelli-Piperno, A. A. & Reich, E. *J. exp. Med.* **159**, 647-652 (1984).
18. Klein, J. R., Pasternack, M. S. & Bevan, M. J. *Immun.* **134**, 2456-2461 (1985).
19. Pasternack, M. S., Sitkovsky, M. V. & Eisen, H. N. *J. Immun.* **131**, 2477-2483 (1983).
20. Laemmli, U. K. *Nature new Biol.* **227**, 680-684 (1970).
21. Huynh, T. V., Young, R. A. & Davis, R. W. in *DNA Cloning Techniques: A Practical Approach* (ed. Glover, D.) 49-78 (IRL Press, Oxford, 1985).
22. Kranz, D. M., Sherman, D. H., Sitkovsky, M. V., Pasternack, M. S. & Eisen, H. N. *Proc. natn. Acad. Sci. U.S.A.* **81**, 573-577 (1984).
23. Reilly, E. B., Kranz, D. M., Tonegawa, S. & Eisen, H. N. *Nature* **321**, 878-880 (1986).
24. Borregaard, N., Heiple, J. M., Simons, E. R. & Clarke, R. A. *J. Cell Biol.* **97**, 52-61 (1983).

Interferon response sequence potentiates activity of an enhancer in the promoter region of a mouse *H-2* gene

Alain Israel, Akinori Kimura, Agnès Fournier*, Marc Fellous* & Philippe Kourilsky

Unité de Biologie Moléculaire du Gène, U277 INSERM, UA 535 CNRS et *Unité d'Immunogénétique Humaine, U276 INSERM, Institut Pasteur, 25 rue du Dr Roux, 75724 Paris Cedex 15, France

The expression of class I transplantation antigens encoded in the major histocompatibility complex (*H-2* in mouse, *HLA* in man) can be induced by α -, β - and γ -interferons¹⁻⁵. Both transcriptional and post-transcriptional mechanisms have been postulated. Recently, a common sequence has been found in the promoter region of several human genes responsive to IFN- α (ref. 6). The promoters of *H-2K^b* and several other mouse class I genes contain a similar interferon response sequence⁷. We show here, in a transient assay, that the *H-2K^b* promoter can be induced by all three types of interferon and that the interferon response sequence is necessary for induction to occur. However, the response sequence is active only when associated with a functional enhancer sequence which we have recently identified in the promoter of *H-2K^b* and other class I genes⁷. The combination of these two sequences can render a heterologous promoter responsive to interferon, irrespective of its orientation relative to the cap site.

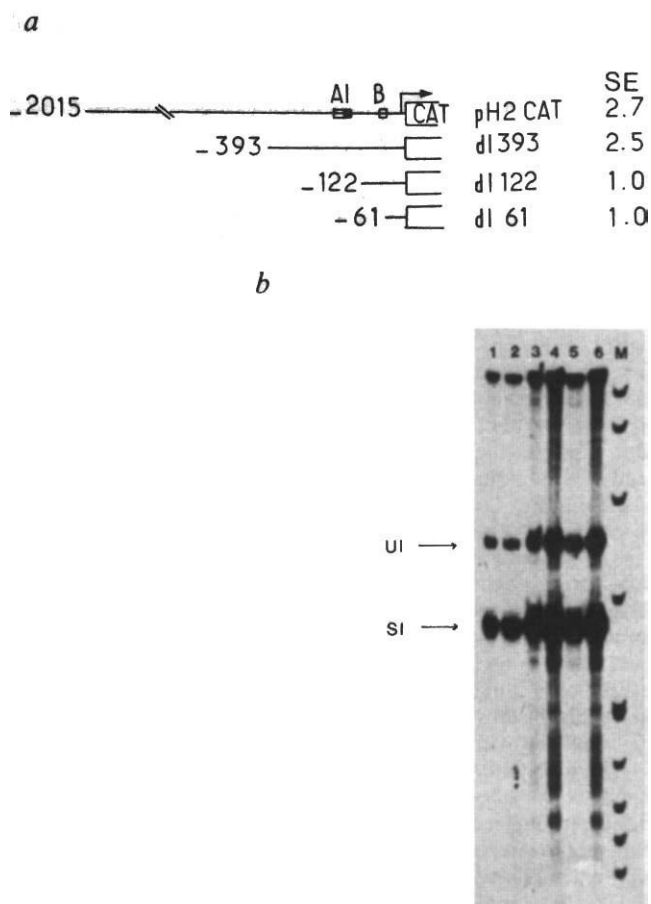
We have previously noted the presence of an interferon response sequence (IRS) in the promoter of several *H-2* class I genes⁷. In *H-2K^b* and *H-2L^d*, the IRS overlaps with an enhancer⁷. To establish the role of IRS in the stimulation of class I gene expression by interferon (IFN), we have used a hybrid gene in which the bacterial chloramphenicol acetyltransferase (CAT) gene is fused to the *H-2K^b* promoter⁸. This construct, called pH2CAT, has been transfected into mouse 3T6 cells. At 16-18 hours after transfection, the cells were treated, or not treated, with mouse type I IFN (100 U ml⁻¹) for 24 hours and CAT activity was assayed. In the same conditions, we confirmed by Northern blotting analysis that transcription of endogenous class I genes is stimulated two- to threefold by type I IFN (not shown). Results in Fig. 1a show that, in this transient assay, CAT activity is increased by IFN treatment. We then tested 5' deletion mutants of pH2CAT and found that *dl393*, which carries the IRS, responds to IFN, while *dl122* and *dl61*, which do not have the IRS, are not stimulated (Fig. 1a). Similarly, the thymidine kinase (TK) and simian virus (SV40) early promoters in the constructs pTKCAT⁷ and pSV2CAT⁹ respectively were not stimulated (not shown). We analysed the

Received 4 February; accepted 22 May 1986.

1. Podack, E. R. & Konigsberg, P. J. *J. exp. Med.* **160**, 695-710 (1984).
2. Henkart, P. A., Millard, P. J., Reynolds, C. W. & Henkart, M. P. *J. exp. Med.* **160**, 75-93 (1984).
3. Podack, E. R., Young, J. D. & Cohn, Z. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8629-8633 (1985).
4. Tite, J. P. & Janeway, C. A. *Eur. J. Immun.* **14**, 878-886 (1984).
5. Tite, J. P., Powell, M. B. & Ruddle, N. H. *J. Immun.* **135**, 25-33 (1985).
6. Nakamura, M., Ross, D. T., Briner, T. J. & Geftter, M. L. *J. Immun.* **136**, 44-47 (1986).

Fig. 1 a, Effect of type I IFN treatment on the activity of pH2CAT and 5' deletion mutants. A (-193 to -159) and B (-76 to -66) indicate the enhancer regions of the *H-2K^b* promoter⁷. I indicates the IFN response sequence (IRS: -165 to -137) which is closely homologous to the consensus sequence described by Friedman and Stark⁶ (see Table 1A). SE, stimulation effect. **b**, Effect of IFN treatment on the level of mRNA induced by pH2CAT and its 5' deletion mutants. Odd-numbered lanes: without IFN; even-numbered lanes: with IFN. Lanes 1, 2, *dl122*; lanes 3, 4, *dl393*; lanes 5, 6, pH2CAT; lane M, markers are pBR322 cut with *Hpa*II and labelled with the Klenow enzyme. The sizes of the specific initiation (SI) and upstream initiation (UI) protected fragments are 283 and 344 nucleotides respectively.

Methods. **a**, Details of the construction of the plasmids and analysis of the enhancer activity have been described before⁷. Duplicate cultures of 3T6 cells (3×10^5 cells per 5-cm dish) were treated for 16–18 h with 0.5 ml of a 1.0-ml calcium phosphate co-precipitate containing 10 μ g of plasmid DNA. The cells were then rinsed once and re-fed with fresh Dulbecco's modified Eagle's medium with 10% fetal calf serum with or without 100 U ml⁻¹ of type I IFN. Mouse type I IFN was a purified preparation (a mixture of IFN- α and - β) given by Drs I. Gresser and M. Tovey (specific activity 2.4×10^7 U mg⁻¹) and used in all experiments. A cytoplasmic extract was prepared 24 h later and CAT activity measured⁷. The protein concentration in the extract was measured with the BioRad protein assay. The ratio of the specific CAT activity (CAT activity/protein concentration) of cells treated with IFN to that of non-treated cells is shown as the stimulation effect (SE). The basal (unstimulated) levels of CAT activity were: 0.45, 1.45, 0.16 and 0.05 units per mg protein for pH2CAT, *dl393*, *dl122* and *dl61* respectively (1 unit represents 1 nanomol of chloramphenicol acetylated per h at 37 °C). These values are the results of both the deletions in the promoter and variable levels of upstream initiation. Experiments were performed at least five times and the average was obtained. Variations in stimulation effect between individual experiments never exceeded 20%. **b**, The steady-state level of CAT mRNA was measured by S₁ mapping analysis. The probe was a uniformly labelled *Eco*RI–*Hinc*II 600-bp fragment derived from *dl61* and cloned in M13mp9; this fragment contains part of CAT gene, 75 bp of the *K^b* promoter (-63 to +12) and part of vector DNA. Using this probe, both SI and UI could be detected⁷. Total RNA was extracted 40 h after transfection from cells treated as described for **a**. To correct for variations in the efficiency of transfection, we included the plasmid pCH110¹⁹ in the transfection mixture and measured the β -galactosidase activity as described previously⁷. An average of 30 μ g of RNA was hybridized with the probe for 16 h at 42 °C. Details for the probe, normalization of the quantity of RNA and S₁ mapping procedure have been described previously⁷.



RNA transcripts by quantitative S₁ mapping experiments, with similar conclusions (Fig. 1b). It is important to note that transcription in pH2CAT and *dl393* is stimulated from the authentic cap site of the *H-2K^b* gene. However, nonspecific initiation from upstream site(s)⁷ located in the vector is also stimulated (Fig. 1b), as if the IRS would act as, or in conjunction with, an

Table 1 Stimulation by cloned human IFNs of *H-2K^b*-derived constructs transfected into HeLa cells

	α	β	γ	A*	IRS*	B*
<i>dl393</i>	1.9	1.8	2.5	+	+	+
<i>dl122</i>	1.0	1.0	1.0	—	—	+
<i>Dde</i> -cII	2.8	3.1	3.9	+	+	—
<i>Dde</i> -Hf	1.0	1.0	1.0	+	—	—

HeLa cells were transfected as described for 3T6 cells (see Fig. 1 legend). IFN- α , - β or - γ were used at 200 units per ml of medium. The values given indicate the stimulation effect due to IFN and are the average of three experiments. The basal CAT values were similar to those obtained with 3T6 cells. Recombinant human IFN- α , - β and - γ have been produced by CHO cells transfected with the corresponding complementary DNAs, and purified by affinity chromatography^{16–18}. Their specific activities were respectively 2×10^8 U mg⁻¹, 5×10^8 U mg⁻¹ and 4×10^8 U mg⁻¹.

* These columns indicate the presence or absence in the construct of enhancer A, interferon response sequence (IRS) and enhancer B, respectively.

enhancer sequence. This observation prompted us to examine more closely the functional relationship between the IRS and the *H-2K^b* enhancers.

We showed previously that two sequences derived from the *H-2K^b* promoter, shown as A and B in Fig. 2a, display an enhancer-like activity when studied in the ponaCAT system⁷. Various fragments cloned in this vector were assayed for stimulation by IFN as described above and the results are shown in Fig. 2a. Stimulation is observed with the *Sau*3A–*Sau*3A fragment (which contains A, B and IRS) in both orientations, with the *Dde*I–*Hinc*II fragment (which contains A and IRS) but not with *Dde*I–*Hinc*II (containing A alone) or *Hinc*II–*Sau*3A (containing IRS and B). We conclude that both enhancer A and IRS are necessary for inducibility, and that they can act in both orientations on a heterologous promoter. Thus, IRS increases the activity of enhancer A upon IFN treatment. Not all enhancers can act jointly with IRS: enhancer B does not, nor does the SV40 enhancer which we cloned upstream of the *Hinc*II–*Sau*3A fragment (Fig. 2a).

The *Q10* gene¹⁰, which maps in the Qa region and is expressed specifically in the liver, also contains an IRS and an enhancer A-like sequence⁷. However, no stimulation by IFN can be observed with either the *Hinc*II–*Hinc*II (which contains the enhancer A-like sequence) or the *Hpa*II–*Hae*III (A-like sequence + IRS) fragment of the *Q10* promoter cloned in ponaCAT (Fig. 2b). Both *Q10* IRS and enhancer A-like sequence display some nucleotide differences from their homologues in *H-2K^b* (Table 2A), and could thus be inactive. In fact, the *Q10* enhancer

Table 2 Sequence of the enhancer A-IRS region in the various constructs described in the text

	SV40 ENHANCER	CONSENSUS IRS	RESPONSE TO IFN
A	CAGGTGTGGAAGTCCCCAGGCT	TTNCNACCTCNGCAGTTTCTCTTCT-CT	
I.	CAGGGGTGGGAAGCCCAGGGCTGGGGATTCCCATCTCCACAGTTTCACTTCTGCA		+
II.	A.G....GG.TCC.ATTGAGAAAC		-
III.	AC.AT.C.TCCGGCGTAGA..A.CC.....		-
IV.	T.....A...C.....C.....T		-
B	ENHANCER A	Q10 IRS	
I.	CAGGGCTGGGGATTCCCAAGCTCCGGATCTCCCATCTCCCAAGTTTCACTTCTGCT		+
II.	A...C.....		-
III.	G.T....AG....G.GCTGA.GA.....		-
IV.	TCA.A.ACACC.GAAAAGGTTGT.A.....		-
A	I : pH2CAT, pd1393, Sau3A-Sau3AconCAT, Dd-cIIconCAT		
	II : Dde-HfconCAT		
	III : Hf-Sau3AconCAT		
	IV : HpaII-Hae(Q10)conCAT		
B	I : Q10 CAT <u>d1167</u> xHf-Hf(K^b), Dde-Hf(K^b), Hf-Hf(L^d)		
	II : Q10 CAT <u>d1167</u> xHf-Hf(Q10)		
	III : Q10 CAT <u>d1375</u>		
	IV : Q10 CAT <u>d1167</u>		

Part of the SV40 enhancer surrounding the core sequence¹⁵ and the consensus IRS⁶ are shown on the first line. Enhancer A of K^b is underlined in AI. The equivalent sequence and the IRS in the Q10 gene are underlined in BI. Lines II, III and IV show the nucleotides differing from line I in the different constructs indicated below. The + sign indicates positive response to IFN, while - denotes no effect (detailed in Figs 1-3).

A-like sequence is at least five times less active than the corresponding sequence of the K^b or L^d genes when assayed in 3T6 cells⁷. We constructed a hybrid Q10-CAT gene and derived two internal deletion mutants, pQ10CAT *d1375* and *d1167* (Fig. 3). We then built hybrid promoters by cloning enhancer A of the K^b , L^d or Q10 genes into pQ10CAT *d1167* (Fig. 3). Neither deletion mutants respond to IFN, but the K^b and L^d sequences restore the responsiveness of the promoter to IFN treatment while Q10 enhancer A-like sequence does not (Fig. 3). In constructs which display IFN induction, specific as well as non-specific initiation is stimulated (confirmed by quantitative S,

mapping analysis, data not shown), similar to results with the K^b promoter (Fig. 1b). Therefore, the IRS of the Q10 gene is potentially active, that is, it can also increase the activity of an enhancer A from K^b or L^d .

Some of these experiments have been repeated with a highly purified mouse IFN- γ preparation (kindly provided by Dr J. Wietzerbin), showing that the IRS sequence also responds to IFN- γ (data not shown). However, since the observed stimulations with mouse type I as well as IFN- γ were only two- to threefold, and since the type I IFN preparation which we used had not been purified to homogeneity, we wished to confirm

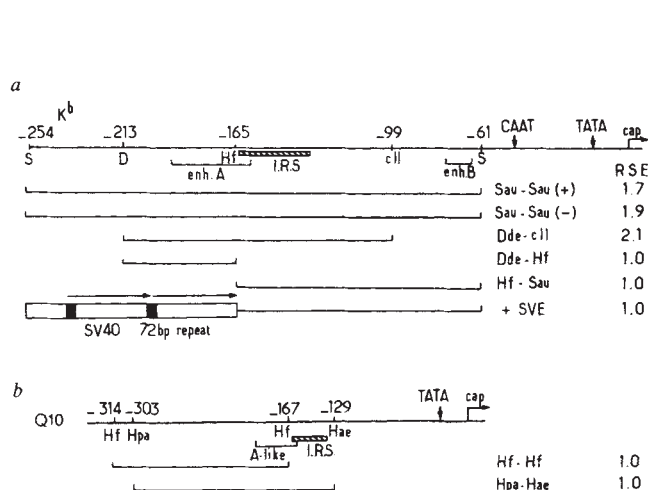
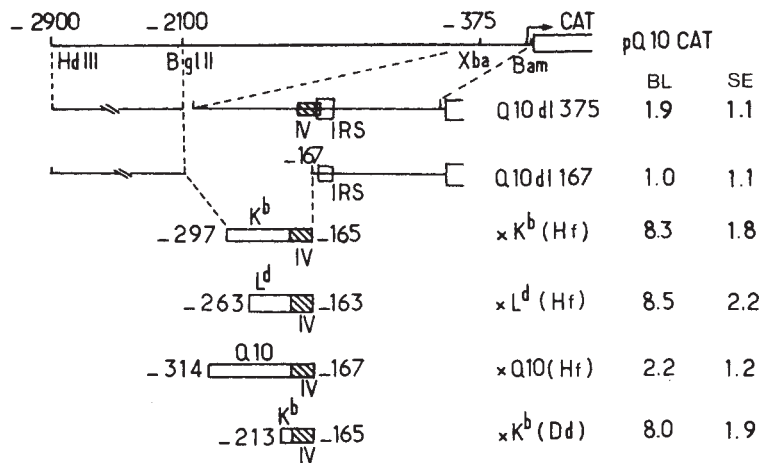


Fig. 2 a, Map of the promoter of $H-2K^b$ and various subfragments cloned in pconCAT. S, *Sau3A*; D, *DdeI*; cII, *HincII*; Hf, *HinfI*; SVE, SV40 enhancer; Hpa, *HpaII*; Hae, *HaeIII*. RSE, relative stimulation effect. Orientation is indicated by arrows and the 'core sequence' by filled boxes. b, Map of Q10 promoter and subfragments cloned in pconCAT.

Methods. Various fragments were cloned at the *Bam*HI site (-102) of pconCAT, in which the CAT gene is driven by the conalbumin promoter (-102 to +62)⁷, using *Bam*HI linkers, except for the *DdeI*-*HinfI* fragment which was cloned by successive addition of *HindIII* and *Bam*HI linkers. Further details for construction of pconCAT and cloning of these subfragments are given in ref. 7. The 72-base-pair (bp) repeats of SV40 were derived from pSV23-2Agpt Δ LR (T. Kadesch, personal communication) as a 200-bp *Bam*HI fragment (coordinates 273 to 95 in ref. 20) and cloned at the 5' *Bam*HI site of the (*HinfI*-*Sau3A*) conCAT construct. Transfection and enzyme assays were performed as described in Fig. 1a legend. Stimulation effect by IFN treatment for each plasmid is indicated as the relative stimulation effect (RSE), being compared with that for the pconCAT construct, which does not respond significantly to IFN. The basal levels of CAT activity in a were 0.97, 0.45, 0.90, 2.2, 0.38 and 8.5 units per mg of protein for the Sau-Sau(+), Sau-Sau(-), Dde-cII, Dde-Hf, Hf-Sau and +SVE constructs, respectively. In b, the basal level of CAT activity was 0.38 units per mg protein for the two constructs.

Fig. 3 Map of the promoter of the *Q10* gene and the *Q10*-CAT hybrid genes. *Xba*, *Xba*I; *Bam*, *Bam*HI; *Hd*III, *Hind*III; (*Hf*), *Hinf*I-*Hinf*I fragment; (*Dd*), *Dde*I-*Hinf*I fragment; IV, sequence identified in ref. 7 which corresponds to enhancer A for *K^b* or *L^d*, and to the enhancer A-like sequence for *Q10*. BL indicates the basal level relative to pQ10CAT *dl*167. The basal CAT activity of this construct was 0.2 units per mg protein.

Methods. A *Hind*III-*Bam*HI fragment (-2,839 to +14) of the *Q10* gene (gene 8 in the cosmid library of Steinmetz *et al.*²¹) was isolated from cosmid 36.2 and cloned with *Hind*III linkers into pCαCAT²² between *Hind*III sites. This construct, named pQ10CAT, was linearized at the unique *Xba*I site and treated with *Bal*31 followed by digestion with *Bgl*II, treatment with Klenow polymerase and re-circularization, to construct internal deletion mutants (pQ10CAT *dl*375 and pQ10CAT *dl*167). The deletion end points were determined by sequencing analysis. A *Hinf*I fragment which contains enhancer A of the *H-2K^b* gene or its equivalent region in the *H-2L^d* or *Q10* gene as well as a *Dde*I-*Hinf*I fragment derived from the *H-2K^b* promoter (*K^b* (*Dd*); see Fig. 2a) were cloned in pQ10CAT *dl*167 at the unique *Bgl*II site (-167) using successive addition of *Hind*III and *Bam*HI linkers (to conserve the sequence of enhancer A 3' to the *Hinf*I site except at one position, see Table 2). These fragments have been shown to have an enhancer activity when assayed in the pconCAT vector system. A *Hinf*I-*Hinf*I fragment derived from the *Q10* promoter has an enhancer activity about five times less than the *K^b* or *L^d* derived fragments⁷. The stimulation effect (SE) of IFN treatment was determined as described in Fig. 1a legend.



our results with recombinant IFN. Mouse recombinant IFN were not easily available to us. We therefore turned to human recombinant IFN. Since *H-2* class I genes transfected into human cells can be regulated by human interferons¹¹, we decided to analyse the regulation of the *H-2K^b* promoter in human HeLa cells, in response to recombinant human IFN- α , - β or - γ . Results in Table 1, obtained with the *dl*122, *dl*393, *Dde*-*Hf* and *Dde*-*cII* constructs, are similar to those obtained in 3T6 cells with mouse IFN, and demonstrate that all three IFNs stimulate the *H-2K^b* promoter in a similar fashion. This strongly suggests that although the receptors for IFN- β and IFN- γ have been shown to be distinct¹², the final target on the *H-2K^b* promoter is the same.

Our experiments with mouse type I IFN demonstrate that the IRS potentiates the action of a functional enhancer in the presence of IFN. This process is specific in the sense that the enhancer A needs to be functional, and that not all enhancers can be made responsive, as illustrated by enhancer B and the SV40 enhancer. The sequence arrangements of enhancer A and IRS in the various constructs described above are shown in Table 2. Interestingly, three mismatches in the enhancer A sequence (Table 2, compare AI and AIV, also BI and BII) are enough to abolish inducibility. Besides, as can be seen in Table 2 (compare AI and BI), enhancer A and IRS can be separated by several nucleotides and still respond to IFN. This suggests that the specific requirement for enhancer A may not be of topological nature. However, we cannot rule out the possibility that the SV40 and B enhancers could be potentiated at an appropriate distance. We have not shown that sequences located in the -137 to -99 region of *H-2K^b* are irrelevant to IFN inducibility, but they are not conserved in the various IFN inducible genes⁶ and we did not find that they have any significant role in class I gene expression⁷.

This functional cooperation between an enhancer sequence and a second regulatory element is reminiscent of the response of the human metallothionein IIA (*MTIIA*) gene to heavy metals¹³. In the present case, the results could be interpreted in terms of a cooperative interaction between putative *trans*-acting factors binding to enhancer A and IRS.

Our results confirm that regulation of class I gene expression

by IFN- α , - β and - γ is, at least partly, transcriptional. We do not exclude additional post-transcriptional mechanisms, considering the relatively low degrees of induction observed and the findings of Yoshie *et al.*¹⁴ that an integrated truncated *HLA* gene, lacking the promoter region, still responds to IFN. Finally, it should be noted that, in the human *MTIIA* gene and in *H-2* and *HLA* class II genes, the IRS is located much further upstream from the cap site than in class I genes⁶. No homology with the class I enhancer A can be found in its vicinity. The mechanism by which these genes are regulated may then differ from class I genes. In this respect, it is noteworthy that maximum induction of the *MTIIA* gene requires 100 times more IFN- α than class I genes⁶.

We thank Drs I. Gresser and M. Tovey for the gift of type I IFN, Dr J. Wietzerbin for purified mouse IFN- γ , Dr D. Novick and Professor M. Revel for human IFN- α , - β , - γ , and Dr M. Steinmetz for making available to us the cosmid clone carrying the *Q10* gene. A.K. was a recipient of an IARC (WHO) fellowship, then a Foundation pour la Recherche Médicale Française fellowship.

Received 23 December 1985; accepted 23 June 1986.

- Lindahl, P., Leary, P. & Gresser, I. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2785-2789 (1973).
- Burrone, O. R. & Milstein, C. *EMBO J.* **1**, 345-349 (1982).
- Wallach, D., Fellous, M. & Revel, M. *Nature* **299**, 833-836 (1982).
- Fellous, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 3082-3086 (1982).
- Rosa, F., Berissi, H., Weissenbach, J., Maroteaux, L., Fellous, M. & Revel, M. *EMBO J.* **2**, 239-243 (1983).
- Friedman, R. L. & Stark, G. R. *Nature* **314**, 637-639 (1985).
- Kimura, A., Israël, A., Le Bail, O. & Kourilsky, P. *Cell* **44**, 261-272 (1986).
- Daniel-Vedele, F., Israël, A., Bénicourt, C. & Kourilsky, P. *Immunogenetics* **21**, 601-611 (1985).
- Gorman, C. M., Moffat, L. F. & Howard, B. H. *Molec. cell. Biol.* **2**, 1044-1051 (1982).
- Devlin, J. J., Lew, A. M., Flavell, R. A. & Coligan, J. E. *EMBO J.* **4**, 369-374 (1985).
- Rosa, F., Hatai, D., Abadie, A. & Fellous, M. *Ann. Inst. Pasteur, Immun.* **136**, 103-119 (1985).
- Branca, A. A. & Baglioni, C. *Nature* **294**, 768-770 (1981).
- Haslinger, A., Heguy, A., Deitman, T., Cooke, T. & Karin, M. (in preparation).
- Yoshie, O. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 649-653 (1984).
- Weier, H., König, M. & Gruss, P. *Science* **219**, 626-631 (1983).
- Novick, D., Eshhar, Z. & Rubinstein, M. *J. Immun.* **129**, 2244-2247 (1982).
- Chernajovsky, Y. *et al. DNA* **3**, 297-308 (1984).
- Novick, D., Eshhar, Z., Fisher, D. G., Friedlander, J. & Rubinstein, M. *EMBO J.* **2**, 1527-1530 (1983).
- Hall, C., Jacob, E., Ringold, G. & Lee, F. J. *molec. appl. Genet.* **2**, 101-109 (1983).
- Fromm, M. & Berg, P. *J. molec. appl. Genet.* **1**, 457-481 (1982).
- Steinmetz, M., Winoto, A., Minard, K. & Hood, L. *Cell* **28**, 489-498 (1982).
- Herbomel, P., Bourachot, B. & Yaniv, M. *Cell* **39**, 653-662 (1984).

Continuous and discontinuous protein antigenic determinants

D. J. Barlow, M. S. Edwards & J. M. Thornton

Laboratory of Molecular Biology, Department of Crystallography, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK

Protein antigenic determinants have been classified as continuous or discontinuous^{1,2}. The continuous determinants are composed of residues which are local in the polypeptide sequence, while discontinuous determinants consist of residues from different parts of the sequence, brought together by the folding of the protein to its native structure. Searches made for protein determinants using peptide fragments which compete with protein-antibody complex formation, or peptides that can be used to raise antibodies which crossreact with the native protein, are limited to the simulation of continuous determinants. However, recent experiments^{2,3} suggest that most determinants are discontinuous. We now show, by consideration of protein surfaces, that if the recognition zone between a protein and antibody has the same dimensions as those found for the lysozyme-antibody complex⁴, none of the protein's surface will be 'continuous'. We suggest that all determinants are discontinuous to some extent, and that crossreacting peptides mimic only the 'primary' interaction site. In addition, we show that the parts of a protein's surface which are most continuous fall predominantly in the loops and/or protruding regions. This explains why quantities such as hydrophilicity⁵, accessibility⁶, mobility⁷ and protrusion⁸ can be used to predict which parts of a polypeptide provide the 'best' antigenic peptides.

The surface of a globular protein, as defined by X-ray crystallographic coordinates, is very complex and convoluted. However, visual inspection of any structure shows that most residues on the surface have neighbouring residues that are distant in the sequence. To quantify this observation, we have used the atomic coordinates available from the protein databank⁹ to search for 'continuous patches' on the surface of a protein. The method used (see Fig. 1a) involves centering a sphere of radius (r) on each surface atom in the protein, and calculating the proportion (F) of the other surface atoms enclosed by the sphere which belong to residues local in the amino-acid sequence. If a sphere encloses only local surface atoms, then we have identified a continuous patch. Surface atoms are defined as those with contact areas (calculated by the method of Lee and Richards¹⁰) $>2\text{\AA}^2$; local surface atoms are those belonging to residues in the sequence $i-n \rightarrow i+n$, where i is the residue containing the surface atom at the origin of a sphere, and n is an integer in the range 1-10.

When the percentage of surface atoms which lie at the centre of a continuous patch is plotted as a function of sphere radius, we obtain the surface 'continuo-grams' shown in Fig. 1b. Four separate plots are presented to show how the results depend on the definition for a local surface atom. These graphs are averaged over three proteins (myoglobin, trypsin and lysozyme), but the individual plots are essentially identical. Note that as the size of the sampling sphere increases, the percentage of the surface which is continuous decreases rapidly, until at a radius of $\sim 10\text{\AA}$ virtually all of the surface is discontinuous, that is, there is no region of $20\text{\AA}$ diameter which contains only atoms from residues local in the amino-acid sequence. This relates directly to the observations made for the lysozyme-antibody complex, where the recognition zone is estimated to be $20 \times 25\text{\AA}$, and is discontinuous⁴. Note however, that this is a monoclonal antibody and has a reasonable affinity for its antigen; the association constant, K_A , is in the range $2-4 \times 10^7\text{ mol}^{-1}$ (R. Poljak, personal communication). Other antibodies that bind more weakly to their antigens may have smaller combining sites, which might there-

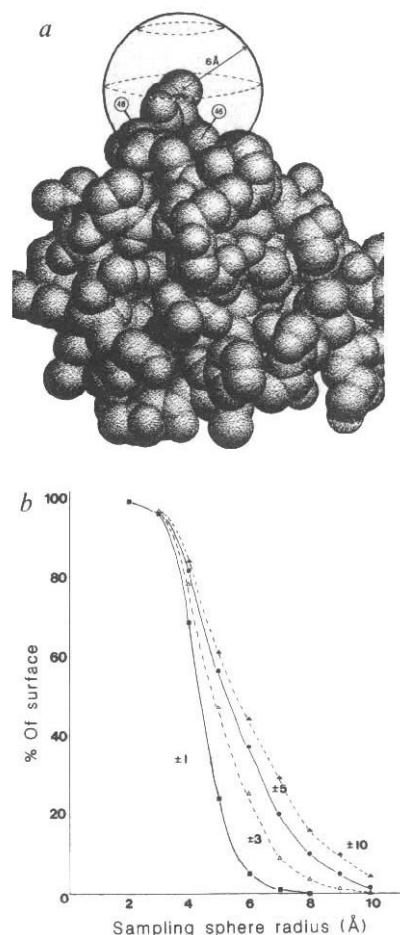


Fig. 1 *a*, Illustration of the method used to calculate the percentage of 'continuous surface' for a protein. The figure shows part of a space-filling model of lysozyme. A sampling sphere of radius $6\text{\AA}$ is centred on one of the surface atoms in the molecule: CG2 of Thr 47. The sphere encloses the atom centres of six other surface atoms, all of which are considered here as local surface atoms, since they belong to residues in the sequence 44-50. Two of the local surface atoms are labelled by residue number. Because all of the atoms enclosed by the sphere belong to local amino-acid residues, CG2 of Thr 47 is considered as the centre of a surface 'continuous patch'. *b*, Surface 'continuo-grams', showing the percentage of surface atoms which lie at the centre of a continuous patch as a function of sampling sphere radius. The four curves are obtained using different definitions for a local surface atom. (Specifically, these are atoms which belong to residues $i-1 \rightarrow i+1$, $i-3 \rightarrow i+3$, $i-5 \rightarrow i+5$, $i-10 \rightarrow i+10$, where i is the residue containing the surface atom at the origin of the sampling sphere.) Note that since the calculations (described in *a* and in the text) consider only atoms whose centres are enclosed by a sampling sphere, these curves represent conservative estimates of the percentage of continuous surface. If the intrusion of any fraction of an atom into the sampling sphere is considered, all of the curves will be shifted to the left.

fore be continuous. As we demonstrate below, however, this is not likely to be the case.

A low-affinity antibody, with $K_A = 10^4\text{ mol}^{-1}$, will have a standard free energy of binding of -6 kcal mol^{-1} (ref. 11). Assuming, therefore, that each $\text{\AA}^2$ of the protein surface buried during association contributes 20 cal (ref. 12) to the binding energy, the buried area will be $\sim 300\text{\AA}^2$. Using sampling spheres of various sizes centred on each of the surface atoms in lysozyme, we calculate that this amount of the protein's surface area would typically involve a patch of $\sim 8\text{\AA}$ radius. As shown in Fig. 1b, $<10\%$ of the surface patches of this size are continuous. Thus,

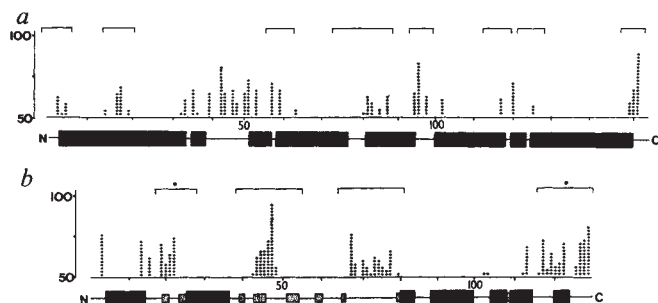


Fig. 2 Variation in $\langle F \rangle$ along the sequence of myoglobin (a) and lysozyme (b). $\langle F \rangle$ values for residues are plotted along the ordinate and the sequences of each protein along the abscissa. The locations of secondary structures in the proteins are shown below the corresponding plot; helices are represented by solid bars, β -strands by stippled bars and loop/turn regions by single lines. The locations of antigenic peptides¹³⁻¹⁵ are indicated above each plot; those labelled in b by • are discontinuous.

binding sites even for low-affinity antibodies are unlikely to involve only local amino-acid residues. We conclude that heptapeptides which cross react with protein antibodies mimic only part of the recognition area. Similarly, anti-peptide antibodies which recognize native protein will be limited to surface patches dominated by a short peptide, where interference from non-local residues is minimal.

On this basis, therefore, it would be reasonable to expect that the parts of a protein that are 'most continuous', should provide the best antigenic peptides. As Fig. 2 shows, this is indeed the case; the two plots illustrate the variation in $\langle F \rangle$ along the sequences of myoglobin and lysozyme, where $\langle F \rangle$ represents the mean value of F for all surface atoms in a residue, calculated using a sampling sphere radius of 10 Å (residues which do not contain surface atoms are assigned a value of $\langle F \rangle = 0$). The 'most continuous' patches in each protein are those where the residues have $\langle F \rangle > 50\%$. Note that there is a strong correlation between these regions and the location of antigenic peptides¹³⁻¹⁵.

An analysis of the surface compositions of 12 proteins shows that ~50% of all loop residues have $\langle F \rangle$ values $> 50\%$ (Fig. 3). The corresponding figures for helix and sheet residues are only 23% and 16%, respectively. As would be expected, therefore, most of the residues which form the most continuous regions lie in loops which protrude from the surface of the protein. Some are also found in helices, but none are found in the central strands of a β -sheet (where residues which are distant in the sequence will always be in close proximity).

These results provide a satisfying rationale, taken together with the necessity for antibody accessibility⁶ or protrusion⁸, to explain which regions of a protein provide the 'best' antigenic peptides.

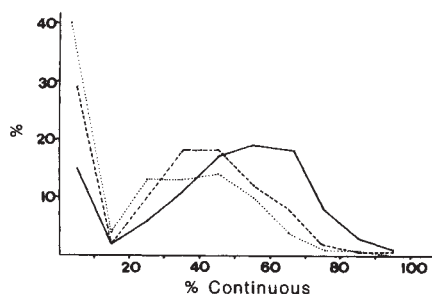


Fig. 3 The distribution of $\langle F \rangle$ values for residues in loop (—), helix (---) and β -strand (·····) regions. % Indicates the frequency of residues with an $\langle F \rangle$ value (% continuous) in a given range.

We thank the SERC for financial support. M.S.E. holds a CASE studentship in collaboration with Dr Peter Murray-Rust at Glaxo.

Received 25 April; accepted 4 June 1986.

1. Atassi, M. Z. *Immunochemistry* **12**, 423-438 (1975).
2. Benjamin, D. C. *et al. A. Rev. Immun.* **2**, 67-101 (1984).
3. Regenmortel, M. H. U. *Hybridoma Technology in Agriculture and Veterinary Research* (eds Stern, N. J. & Cramble, H. R.) Ch. 3 (Rowman & Allanheld, New Jersey, 1984).
4. Amit, A. G., Mariuzza, R. A., Phillips, S. E. V. & Poljak, R. J. *Nature* **313**, 156-158 (1985).
5. Hopp, T. P. & Woods, K. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3824-3828 (1981).
6. Novotny, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 226-230 (1986).
7. Westof, E. *et al. Nature* **311**, 123-126 (1984).
8. Thornton, J. M., Edwards, M. S., Taylor, W. R. & Barlow, D. J. *EMBO J.* **5**, 409-413 (1986).
9. Bernstein, F. C. *et al. J. molec. Biol.* **112**, 535-542 (1977).
10. Lee, B. & Richards, F. M. *J. molec. Biol.* **55**, 379-400 (1971).
11. Schulz, G. E. & Schirmer, R. H. *Principles of Protein Structure* Ch. 3 (Springer, New York, 1979).
12. Richards, F. A. *Rev. Biophys. Bioengng* **6**, 151-176 (1977).
13. Atassi, M. Z. *Eur. J. Biochem.* **145**, 1-20 (1985).
14. Ibrahim, I. M., Eder, J., Prager, E. M., Wilson, A. C. & Arnon, R. *Molec. Immun.* **17**, 37-46 (1980).
15. Takagaki, Y., Hirayama, A., Fujio, H. & Amano, T. *Biochemistry* **19**, 2498-2505 (1980).

Expression of a transfected human c-myc oncogene inhibits differentiation of a mouse erythroleukaemia cell line

Ethan Dmitrovsky*, W. Michael Kuehl*, Gregory F. Hollis*, Ilan R. Kirsch*, Timothy P. Bender* & Shoshana Segal*†

* NCI-Navy Medical Oncology Branch, National Cancer Institute, and † Department of Medicine, Uniformed Services University of the Health Sciences, National Navy Medical Center, Bethesda, Maryland 20814-2015, USA

The Friend-virus-derived¹ mouse erythroleukaemia (MEL) cell lines represent transformed early erythroid precursors that can be induced to differentiate into more mature erythroid cells by a variety of agents including dimethyl sulphoxide (DMSO)². There is a latent period of 12 hours after inducer is added, when 80-90% of the cells become irreversibly committed to the differentiation programme, undergoing several rounds of cell division before permanently ceasing to replicate^{3,4}. After DMSO induction, a biphasic decline in steady-state levels of c-myc^{5,6} and c-myb⁶ messenger RNAs occurs. Following the initial decrease in c-myc mRNA expression, the subsequent increase occurs in, and is restricted to, the G₁ phase of the cell cycle⁷. We sought to determine whether the down-regulation is a necessary step in chemically induced differentiation. Experiments reported here indicate that expression in MEL cells of a transfected human c-myc gene inhibits the terminal differentiation process.

To study the relationship between the down-regulation of the c-myc mRNA and cellular differentiation, a plasmid (PL¹hmcneo¹) containing the pSV2neo and the human c-myc genes was transfected into MEL cells (line 745). The stable G418-resistant transfectants (Fig. 1) were induced with DMSO to determine their differentiation potential. Strikingly, all transfectants which expressed the exogenous human c-myc gene failed to develop a red pellet or a positive reaction with benzidine after 7 days of induction. This inhibition of differentiation is not a transfection artefact. The parental MEL cells were subcloned after undergoing a sham transfection without selection, and all 24 subclones examined differentiated after 7 days of DMSO induction. In addition, transfectant T57, which expressed only the neomycin resistance gene, differentiated normally after induction. Introduction of another selection marker, the APRT (adenine phosphoribosyltransferase) gene, also does not interfere with induced differentiation of MEL cells⁸.

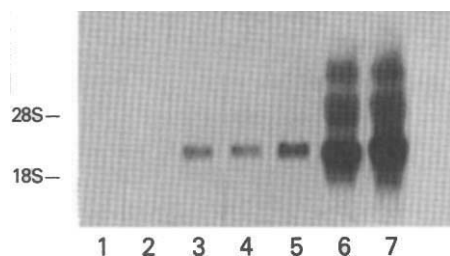


Fig. 1 RNA expression of the exogenous human *c-myc* in stable MEL transfectants. The inducible murine Friend-virus-derived¹ erythroleukaemia cell line 745 (a gift from Dr J. Billelo) was transfected by a modified¹⁶ calcium phosphate technique¹⁷ with PL¹hmcneo⁺ (a gift from Dr W. Lee). This plasmid is a derivative of the pCV108 vector¹⁸. It contains a 5.8-kb *Xho*I/*Eco*RI genomic human *c-myc* fragment (including all three *c-myc* exons) fused at the *Xho*I site of the first exon to a portion of the U3 and R of the Moloney murine leukaemia LTR. Total cellular RNA was prepared from the uninduced MEL parent and stable transfectants using the guanidine thiocyanate method previously described¹⁹. Total cellular RNA (10 µg per lane) was loaded, electrophoresed on a 1% agarose formaldehyde gel and blotted²⁰. The nitrocellulose filter was hybridized to a ³²P-dCTP nick-translated human *c-myc* 1.5-kb *Cl*aI/*Eco*RI third exon probe overnight at 42 °C in a 10% dextran sulphate, 5×SSC, 40% formamide solution. After hybridization, the filter was washed for 1 hour at 60 °C with 0.1×SSC, 0.1% SDS and analysed by autoradiography using Kodak XRP-5 film and an overnight exposure with an intensifying screen. The five transfectants were obtained from three independent transfection experiments. The numbered lanes show the 745 MEL parent (lane 1), T57 (lane 2) T61 (lane 3), T62 (lane 4), T60 (lane 5), T56 (lane 6) and T55 (lane 7).

Two transfectants were selected for further studies to better define the effect of exogenous human *c-myc* expression on DMSO induction: T56, which expressed high levels of human *c-myc* mRNA, and T57 (Fig. 1). Both T56 and T57 were cloned by limiting dilution. Since the usual induction procedure (designated stationary phase induction) is limited by the 3–4 days taken for MEL cells to reach stationary growth^{2,3}, we modified this procedure to continue the induction longer. MEL cells were fed with enough medium (with or without DMSO) each day to maintain logarithmically growing cells at approximately the same density (designated logarithmic phase induction). Exponentially growing cells of clone 56 and 57 were seeded in medium without (control) or with (induced) DMSO. As shown (Fig. 2), clone 57, like the parental MEL cells, accumulated benzidine-positive cells to a maximum level of 90%. In contrast, cells from clone 56 were benzidine unreactive and the cell pellet remained white even after 14 days of induction. Another assay for terminal differentiation is clonability. Following DMSO induction, clone 57 rapidly lost its ability to generate subclones, while clone 56 showed no decrease in clonability after 14 days.

Steady-state levels of α -globin, *c-myc* and β -actin mRNAs, as well as exogenous human *c-myc* mRNA, were determined by Northern blots of total RNA prepared from clones 56 and 57 with or without DMSO (Fig. 3). The expression of β -actin was similar for the two clones, and remained essentially constant for control or induced cultures. Expression of exogenous human *c-myc* was stringently assayed so that the endogenous *c-myc* mRNA was not detected. As shown (Fig. 3), clone 56 expressed relatively high levels of two human *c-myc* mRNA species (approximately 2.3 and 2.1 kilobases (kb) in size). It is possible that transcription of these two mRNA species is initiated at the long terminal repeat (LTR) and P2 *c-myc* promoters⁹, respectively. Notably, the expression of human *c-myc* mRNA decreased 6 hours after DMSO was added, and then declined more gradually during the next 10 days.

Using an α -globin cDNA probe, clone 57 expressed a basal level of α -globin mRNA which increased substantially during

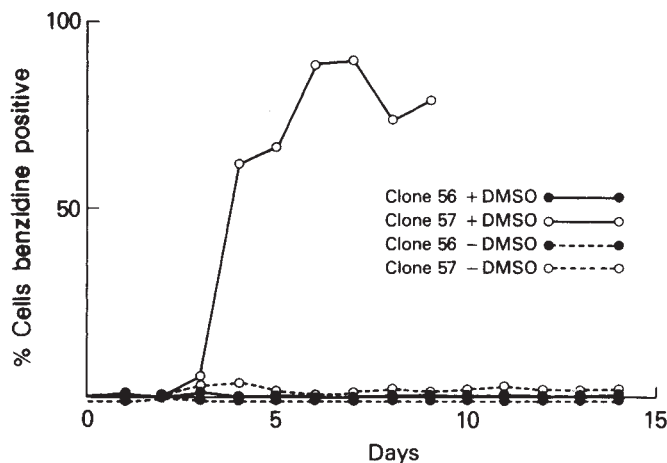


Fig. 2 Accumulation of benzidine-stained cells of clones 56 and 57 for uninduced and 'logarithmic-phase' induced cultures. Logarithmically growing cells of these clones were seeded at a density of 5×10^5 per ml on day -1 and split to 10^5 per ml one day later (day 0) in Joklik's modified medium supplemented with $75 \mu\text{g ml}^{-1}$ of G418 (Geneticin, Gibco) and 10% fetal calf serum with or without 1.5% DMSO (Sigma). Each day half of the cells were removed from culture, harvested by centrifugation, washed twice with phosphate-buffered saline (PBS), and then stained for haemoglobin using the acid benzidine stain as previously described³. The remaining cells were replenished with an equal volume of fresh Joklik's modified medium containing G418 and 10% fetal calf serum, with or without 1.5% DMSO as indicated. For both the induced and uninduced cultures, 200 cells were daily assessed for benzidine reactivity and the percentage of benzidine-positive cells are provided for the indicated time points. The cell pellet of clone 57 became red approximately one day before the majority of individual benzidine-positive cells were seen.

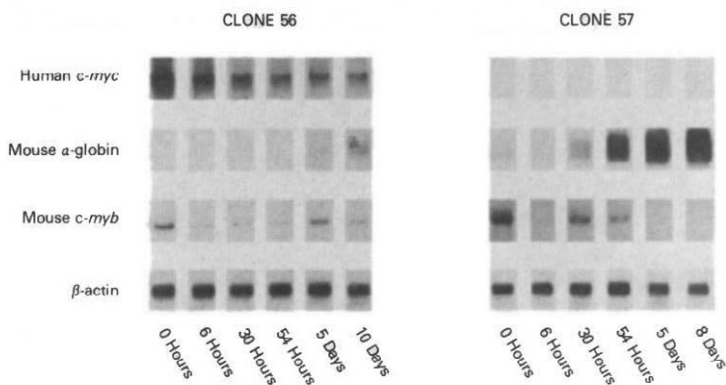
DMSO induction. This increase in α -globin mRNA expression was comparable in kinetics and magnitude to that observed with the parental 745 MEL line under stationary-phase induction conditions^{5,6}. In marked contrast, clone 56 expressed substantially lower basal levels of α -globin mRNA, which increased minimally after 10 days of induction. We do not know if this reflects low levels of expression in all cells or selective expression in a subpopulation of cells.

Changes in the expression of *c-myc* mRNA provide another marker for DMSO response. As expected, for clone 57 the changes in *c-myc* were indistinguishable from those previously observed⁶ for parental MEL cells (Fig. 3). For clone 56, the basal *c-myc* mRNA level was lower than for clone 57. This level of mRNA decreased within 6 hours after induction. Unlike the parental cells or clone 57, the level of *c-myc* mRNA stayed relatively constant during the remainder of the induction period.

To compare changes in expression of endogenous *c-myc* mRNA in the transfected cells, we used an *S*₁ nuclease protection assay to ensure that the endogenous mouse *c-myc* could be distinguished from the exogenous human *c-myc* expressed in clone 56 (Fig. 4). Despite the presence of high levels of exogenous *c-myc* mRNA, the endogenous *c-myc* mRNA continues to be expressed and the initial transient decline occurs although the second decline does not even after 10 days of induction.

The six parameters used to assess the response of clones 56 and 57 to DMSO induction were: cell pellet colour and benzidine staining of single cells; endogenous α -globin, *c-myc* and *c-myc* mRNA levels; and cell cloning efficiencies. These experiments clearly indicate a failure of clone 56 to differentiate normally, unlike clone 57 or the parental MEL cells. This inhibition occurs in five transfectants that expressed various levels of exogenous human *c-myc* mRNA (Fig. 1)¹⁰. These results are consistent with evidence from the avian myoblast cell culture system where myoblasts infected with retrovirus and expressing high levels of

Fig. 3 Expression of human *c-myc*, α -globin, *c-myb* and β -actin mRNAs in clones 56 and 57 during logarithmic-phase induction by DMSO. Logarithmically growing cells of clones 56 and 57 were seeded, grown and harvested as described in Fig. 2. However, after washing in PBS, the removed cells were resuspended in guanidine thiocyanate for total cellular RNA analysis. Total RNA (10 μ g) was applied to each lane, electrophoresed on a 1% agarose/formaldehyde gel and blotted²⁰. The obtained nitrocellulose filters were hybridized as described in Fig. 1 to the following probes: *a*, 1.5-kb *Cla*I/*Eco*RI third exon and 3' flanking human *c-myc* probe; *b*, 3-kb *Sst*I fragment representing a murine α -globin cDNA prepared from a plasmid provided by Dr Y. Nishioka; *c*, 582 base pairs (bp) *Sma*I/*Eco*RI murine *c-myb* cDNA fragment²¹; *d*, 1.9-kb *Bam*HI fragment representing approximately full-length cDNA for human β -actin provided by Dr L. Kedes. The hybridized filters were then washed at 54 °C as described in Fig. 1 except for filters hybridized to the *Cla*I/*Eco*RI human *c-myc* which were washed at 62 °C. Autoradiographs were obtained after exposure to XRP-5 film and an intensifying screen. Steady-state levels of the indicated mRNAs remained relatively constant for uninduced cultures. Densitometric analysis of Southern blots demonstrated 36 copies of the human *c-myc* gene integrated in the genome of clone 56 whereas clone 57 does not contain an intact human *c-myc* gene (data not shown).



v-myc protein fail to differentiate into myotubes¹¹. Recently, similar findings were reported in a thymidine kinase—MEL cell line where transfection of a simian virus 40—driven *c-myc* inhibits erythroleukaemic differentiation¹². Experiments to determine whether other oncogenes can substitute for this inhibitory effect are under way.

Some tumours can be induced to differentiate in culture. In many of these cases, the expression of the *c-myc* and *c-myb* oncogenes decreases during differentiation^{5,6,13–15}. Two patterns of decline have been observed. The first is a gradual disappearance of *c-myc* or *c-myb* mRNAs following induction of differentiation, as seen in the HL60¹³, WEHI3B¹⁴, and F9 cell lines¹⁵. The second is a biphasic decline in steady-state levels of *c-myc*^{5,6} and *c-myb*⁶ mRNA as seen in MEL cells. In general, malignant cells appear 'frozen' in a particular state of differentiation. Since qualitative or quantitative changes in oncogene expression are implicated as a cause of malignant transformation, it is possible that the apparent block in differentiation of tumour cells is due to abnormal oncogene expression. Our results suggest that *c-myc* expression plays a central role in determining whether a cell remains in a continuous cycle of growth and proliferation or enters a differentiation pathway.

We thank William Lee for use of the PL^{hmcneo} plasmid,

James Battey for helpful discussions, and Virginia Bertness and Laura Minna for technical assistance.

Received 31 March; accepted 16 May 1986.

1. Friend, C., Scher, W., Holland, J. G. & Sato, T. *Proc. natn. Acad. Sci. U.S.A.* **68**, 378–382 (1971).
2. Marks, P. A. & Rifkind, R. A. *Rev. Biochem.* **47**, 419–448 (1978).
3. Fibach, E., Reuben, R. C., Rifkind, R. A. & Marks, P. A. *Cancer Res.* **37**, 440–444 (1977).
4. Gusella, J., Geller, R., Clarke, B., Weeks, V. & Housman, D. *Cell* **9**, 221–229 (1976).
5. Lachman, H. M. & Skoultschi, A. I. *Nature* **310**, 592–594 (1984).
6. Kirsch, I. R., Bertness, V., Silver, J. & Hollis, G. F. *J. cell. Biochem.* (in the press).
7. Lachman, H. M., Hatton, K. S., Skoultschi, A. I. & Schildkraut, C. L. *Proc. natn. Acad. Sci. U.S.A.* **82**, 5323–5327 (1985).
8. Charnay, P. *et al. Cell* **38**, 251–263 (1984).
9. Yang, J. Q. *et al. Curr. Topics Microbiol. Immun.* **113**, 146–153 (1984).
10. Dmitrovsky, E. *et al. Curr. Topics Microbiol. Immun.* (in the press).
11. Falcone, G., Tato, F. & Alema, S. *Proc. natn. Acad. Sci. U.S.A.* **82**, 426–430 (1985).
12. Coppola, J. A. & Cole, M. D. *Nature* **320**, 760–763 (1986).
13. Westin, E. H. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 2490–2494 (1982).
14. Gonda, T. J. & Metcalf, D. *Nature* **310**, 249–251 (1984).
15. Campisi, J., Gray, H. E., Pardee, A. B., Dean, M. & Sonenshein, G. E. *Cell* **36**, 241–247 (1984).
16. Davis, L., Dibner, M. & Battey, J. *Methods in Molecular Biology* (Elsevier, Amsterdam, 1986).
17. Graham, F. L. & Van der Eb, A. J. *Virology* **52**, 456–467 (1973).
18. Lau, Y.-F. & Kan, Y. W. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5225–5229 (1983).
19. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294–5299 (1979).
20. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
21. Bender, T. P. & Kuehl, W. M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3204–3208 (1986).
22. Battey, J. *et al. Cell* **34**, 779–787 (1983).

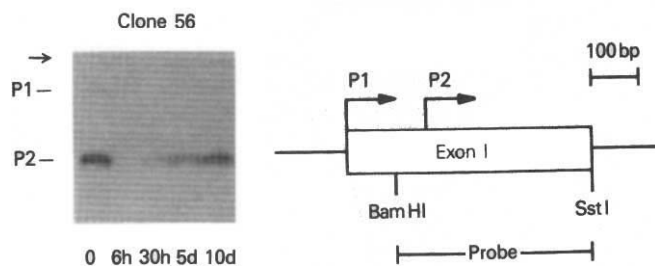


Fig. 4 Expression of endogenous *c-myc* mRNA determined by *S*₁-nuclease protection with clone 56 total cellular RNA obtained during logarithmic-phase induction (from the experiment described in Fig. 3). The single-stranded *S*₁ probe was prepared by primer extension of an M13 clone containing the 400-bp *Bam*HI/*Sst*I fragment derived from exon I of a genomic murine *c-myc* clone (obtained from Dr J. Battey). The probe was hybridized overnight at 55 °C to 10 μ g of RNA using the previously described method²². After digestion with *S*₁-nuclease, the fragments were fractionated on an 8 M urea-5% polyacrylamide gel. The time points are hours (h) or days (d) after initiation of DMSO induction. The arrow indicates the position of the undigested probe. P1 and P2 represent promoters 1 and 2 of murine *c-myc* gene⁹.

DNA loops induced by cooperative binding of λ repressor

Jack Griffith, Ann Hochschild* & Mark Ptashne*

Lineberger Cancer Research Center, University of North Carolina Medical School, Chapel Hill, North Carolina 27514, USA

* Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

It has been shown by Hochschild and Ptashne¹ that λ repressors bind cooperatively to operator sites separated by five or six turns of the helix. Cooperative binding is not observed if the sites are separated by a nonintegral number of turns, unless a four-nucleotide gap is introduced into one of the strands between the two sites. These and other facts suggested that repressors at the separated sites touch each other, the DNA bending smoothly so as to accommodate the protein–protein interaction. Here we use electron microscopy to visualize the predicted protein–DNA complexes.

Figure 1 shows the DNA molecules used in our experiments as well as the anticipated protein–DNA complexes. Fragments I and III bear two operator sites separated by five helical turns,

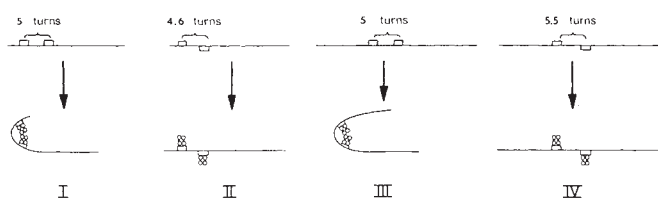


Fig. 1 DNA molecules used and expected shapes of complexes formed with λ repressor. The centre-to-centre distances separating the operator sites are given in helical turns and the molecules are drawn approximately to scale. DNA sequences spanning the two operator sites for each of these molecules are as previously described¹.

Table 1 Electron microscopic count of DNA fragments

DNA fragment	No. straight	No. bent	% Bent
1. I	36	8	19
2. I	50	32	39
3. II	213	0	0
4. II	37	0	0
5. II	32	0	0
6. II	46	1	2
7. III	113	9	7
8. III	67	7	9
9. IV	106	0	0
10. IV	86	0	0
11. IV	38	0	0

Each line represents a separate experiment. The repressor concentration in the first four experiments was one half that of the other experiments.

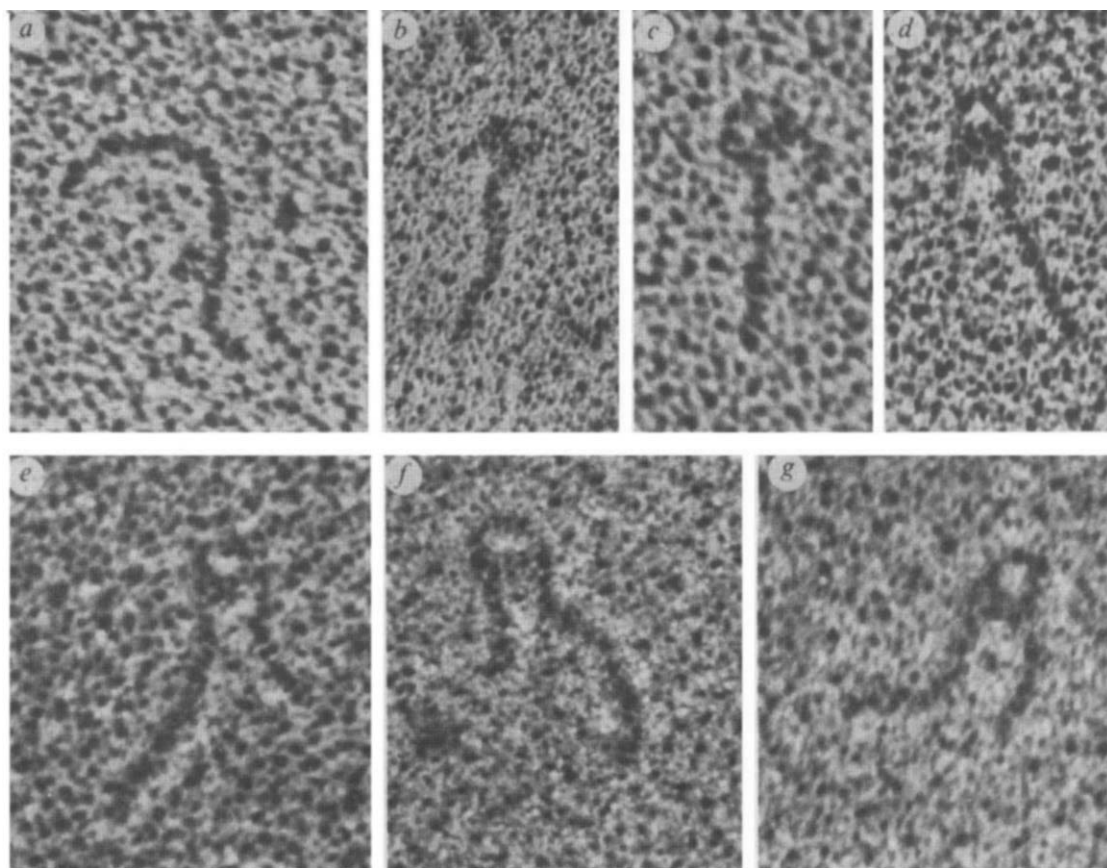


Fig. 2 Visualization of λ repressor bound to DNA fragments shown in Fig. 1. Purified λ repressor is shown bound to DNA fragments I (b-d), III (e-g) and II (a), as visualized by electron microscopy. In these experiments, 300–400 base-pair fragments containing repressor binding sites (20 ng ml^{-1}) were incubated with purified repressor (2 or $4 \mu\text{g ml}^{-1}$) in a buffer containing 10 mM Tris-HCl, $\text{pH } 7.0$; 25 mM MgCl_2 ; 10 mM CaCl_2 ; 1 mM EDTA and 200 mM KCl. Supertwisted pBR322 DNA was included at $2 \mu\text{g ml}^{-1}$ to reduce nonspecific protein binding. Following 20-min incubations at 37° , the complexes were fixed by the addition of HEPES buffer to 20 mM and formaldehyde to 1% for 5 min at 20° , followed by the addition of glutaraldehyde to 0.6% for an additional 5 min . The samples were directly mounted onto thin carbon films, washed with water and ethanol rinses, air dried and rotary shadowcast with tungsten as previously described¹⁰. Shadowcasting was done in a fully cryopumped system at a vacuum of 2×10^{-7} torr. The micrographs were taken with a Phillips EM400 instrument operated at 20 kv . Bar equals $0.1 \mu\text{m}$.

assuming 10.5 base pairs per turn and measuring from the central base pair. The disposition of the two sites with respect to the ends of the molecule are different in the two cases and the protein-DNA complexes were expected to be distinguishable under the electron microscope. Fragments II and IV bear operator sites separated by 4.6 and 5.5 turns respectively and in this case we expected to see no looped DNA structures upon addition of repressor. In each case the two sites were those used previously¹. One is the wild-type site O_R1 and the other a mutant

derivative O_R1^m that binds repressor sixfold more weakly than its wild-type parent. Experiments using DNase I to measure equilibrium binding show that repressor binds cooperatively to the separated sites on molecules I and III but non-cooperatively to the sites on molecules II and IV¹.

Figure 2 shows representative electron micrographs of protein-DNA complexes using three of our four DNA fragments. These samples were prepared by incubating repressor and DNA in conditions like those used to measure cooperative binding

using the DNase footprinting assay¹. The samples were then fixed and shadowed with tungsten (see legend to Fig. 2). At the magnification required to visualize the short DNA fragments, two or more fragments were seldom found in the same field. Nonetheless it was not difficult to count and examine a statistically significant number of molecules for each sample. The DNA fragments were scored visually in the electron microscope as being either straight or bent with a small protein complex on the inside of the bend. Those that were scored as bent were photographed and subsequently examined to confirm this classification. Table 1 summarizes the distributions of molecules observed.

Many examples of DNA fragment I were found in which the DNA was bent into a J shape with protein bound on the inside of the bend. (Fig. 2b-d; Table 1, lines 1 and 2). The loop between the bound repressors measured 15 ± 5 nm, corresponding to 50 ± 15 base pairs. The location of the loop along the length of the 300-base-pair fragment and the length of the short arm (25 ± 5 base pairs) were consistent with the complex drawn in Fig. 1.

In contrast, only one in the 238 molecules of DNA fragment II examined appeared bent when repressor was complexed with this fragment (Fig. 2a; Table 1, lines 4-6). Repressor concentrations tenfold higher than those needed to produce looped structures with fragments I and III also failed to produce such structures (data not shown).

DNA fragments III and IV were prepared in Cambridge and sent to Chapel Hill labelled in code. Numerous bent complexes were found with one DNA sample (Fig. 2e-g; Table 1, lines 7 and 8) but none with the other (Table 1, lines 9-11). As expected the former proved to be fragment III; the latter fragment IV. Note that the size of the loop observed with fragment III is

similar to that observed with fragment I but that the shorter of the two arms emerging from the complex is approximately four times longer. This result is as predicted (see Fig. 1).

For all four fragments examined, it was common to observe what appeared to be a repressor on the DNA at one or the other binding site (see for example Fig. 2a). In most cases, the repressor appeared to be bound to the side of the DNA. As expected¹, there is no indication that the binding of a repressor to a single operator site bends the DNA. Very few fragments were found with two repressors bound, however, even though this was expected with fragments II and IV. It is possible that the fixation steps used here may have released many singly bound repressors from the DNA.

Several examples have now been reported of apparent protein-protein interactions "at a distance" that regulate gene expression²⁻⁸. Some if not all of these cases presumably involve formation of DNA loops^{2,3,9}. The electron micrographs described here provide the first direct visualization of loops formed when proteins bind cooperatively to separated sites on the DNA.

This research was supported in part by NIH grants GM31819 to J.G. and GM22526 to M.P., A.H. was supported in part by NIH Genetics Training grant GM07620 to the Department of Cellular and Developmental Biology, Harvard University.

Received 23 May; accepted 6 June 1986.

1. Hochschild, A. & Ptashne, M. *Cell* **44**, 681-687 (1986).
2. Dunn, T., Hahn, S., Ogden, S. & Schleif, R. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5017-5020 (1984).
3. Majumdar, A. & Adhya, S. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6100-6104 (1984).
4. Dandaneil, G. & Herman, K. *EMBO J.* **4**, 3333-3338 (1985).
5. Takahashi, K. *et al. Nature* **319**, 121-126 (1985).
6. Brent, R. & Ptashne, M. *Cell* **43**, 729-736 (1985).
7. Keegan, L., Gill, G. & Ptashne, M. *Science* **231**, 699-704 (1986).
8. Riddihough, G. & Pelham, H. R. B. *EMBO J.* (in the press).
9. Ptashne, M. *A Genetic Switch* (Cell, Palo Alto & Blackwell, Cambridge 1986).
10. Griffith, J. & Christiansen, C. A. *Rev. Biophys. Bioengng.* **7**, 19-37 (1978).

Length growth in fission yeast cells measured by two novel techniques

J. W. May* & J. M. Mitchison

Department of Zoology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK

Time-lapse films have shown that cells of the fission yeast *Schizosaccharomyces pombe* exhibit a sharp increase in the rate of length growth at one stage of the cell cycle and that the same periodic increase in growth rate, a cell-cycle event, occurs in a cell-cycle mutant after the DNA division cycle has been blocked¹. The continued periodicity of cell-cycle events after such a block has been shown in other systems^{2,3} and has intriguing implications for cell-cycle control models. We have therefore developed two new methods for measuring the rate of length growth which are equivalent to pulse labelling. The results confirm the evidence from films and also show that the rate of growth reaches a plateau after the blocked cells have achieved a critical size.

Time-lapse films are the most direct way of measuring cell growth, but analysis accurate enough for rate determinations is very time-consuming. In addition, some cell cycle mutants only grow under mounting conditions that prevent good optical resolution¹. Our methods made use of the fact that *Bandeiraea simplicifolia* lectin (BSL) binds to the cell wall of *S. pombe*⁴. They also took account of the need to measure cell size as well as rate of growth, since the point where the rate changes in wild-type cells is size-related¹. In our initial experiments with wild-type cells, an overall coating of BSL coupled to fluorescein isothiocyanate (FITC) was applied. The cells were then washed

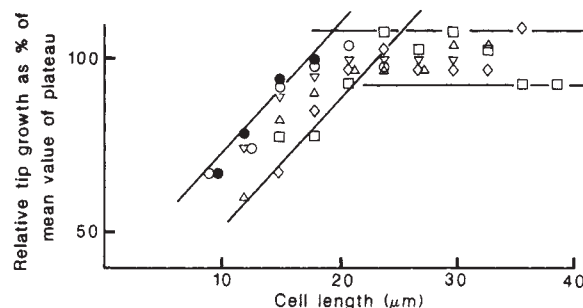


Fig. 1 Relative rate of tip growth against cell length in *S. pombe* *cdc 2.33* at increasing times after shift-up to the restrictive temperature. Cells were grown up overnight at 28 °C in a minimal medium EMM3 (ref. 9) + 0.5% yeast extract + 0.1 $\mu\text{g ml}^{-1}$ biotin (this extra biotin was to guard against binding by traces of avidin). The medium was adjusted to pH 6.0 to promote lectin binding. The culture (1.2×10^6 cells ml^{-1}) was shifted to 36.5 °C at time zero. Samples were taken at 1 h (●), 2 h (○), 2.5 h (▽), 3 h (△), 3.5 h (◇) and 4 h (□). Each sample was treated as follows: 1 min in biotinylated BSL (type BS 1) at 5 $\mu\text{g ml}^{-1}$; washed once; 1 min in avidin conjugated with Texas Red at 1.3 $\mu\text{g ml}^{-1}$ plus avidin at 5 $\mu\text{g ml}^{-1}$; washed twice; suspended in conditioned medium and incubated, one part (experimental) at 36.5 °C for 30 min and the other part (control) at 0 °C; then tip coating put on by 1 min in BSL-FITC at 2 $\mu\text{g ml}^{-1}$; washed once and preserved with formalin at 1%. Washing was done by centrifugation in a small bench centrifuge for 30 s at 13,400g. All reagents were from Sigma. Analysis was done in a FACS IV (Becton Dickinson) on 25,000 cells. The results were grouped for red fluorescence (length) and the peak value found for the green fluorescence (tip growth). There was appreciable penetration of the green lectin into the areas coated by the first red lectin. This was allowed for by taking the peak value of the green fluorescence in the control above (where there had been no tip growth) and subtracting this from the peak value of the experimental sample. The red fluorescence has been converted into cell length by taking the results from an earlier experiment in which cell length was measured against time after shift-up.

* Present address: Department of Microbiology, Monash University, Clayton, Victoria 3168, Australia.

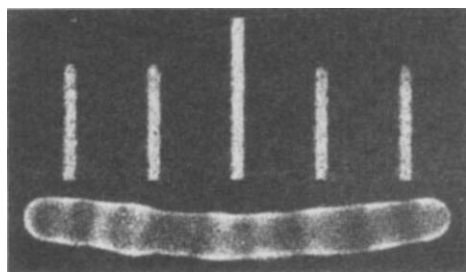


Fig. 2 A cell of *S. pombe cdc 2.33* growing at both ends and coated alternately with lectin with FITC and lectin without FITC; a 'double-ended tiger-tail'. The coating procedure was as for Fig. 3 except that the intervals between coatings was 40 min. Scale divisions at 10- μ m intervals.

and left to grow for 30 min (20% of the generation time). During this time, new uncoated wall was generated at the ends of the cell, as *S. pombe* grows by tip extension. At the end of the period, the new wall was coated with the same lectin but coupled with Texas Red through a biotin-avidin link. Under fluorescence microscopy, the cells were green with red tips. There was an intermediate yellow zone between body and tips due to overlap between the two lectins, and a mixing of old and new wall material at the growing tip.

Most of the experiments were done with a temperature-sensitive mutant of *S. pombe cdc 2.33* (ref. 5). When the culture was shifted to the restrictive temperature (36.5°C), mitosis was blocked but the cells continued to grow for several generation times and became very elongated. The coating procedure described above was used on cells of *cdc 2.33* except that the sequence was reversed, with Texas Red being used as the body coating and FITC as the tip coating. This was done on samples taken every 30 min from 1 h after shift for 4 h. The samples were then analysed in a fluorescence-activated cell sorter (FACS) which measured each cell for red fluorescence (cell length minus tips) and green fluorescence (tip growth). The grouped results in Fig. 1 show that, as would be expected, the later samples had more large cells and fewer small cells. They also show that, irrespective of sampling time, the rate of tip growth increased with cell length up to a limiting length of about 20 μ m. Thereafter the rate plateaued. These results are quite consistent with the rate changes observed in the time-lapse films since sharp steps in rate in individual cells would be spread out to give a slower rise in cell populations. What is missing is an initial rate plateau in the smallest cells before the rise. It proved impossible, however, to obtain reliable measurements of green fluorescence in these cells because they were contaminated by a subpopulation of small non-growing cells.

The second method used a coating procedure similar to the first one except that no fluorochrome was conjugated to the avidin. The important difference was that the procedure was continued for 6 h with alternate treatments every 30 min with lectin plus FITC and with lectin without a fluorochrome. The cells were photographed with a fluorescence microscope at the end of the experiment. They showed a pattern of bands (we call them 'tiger-tails') which define the history of length growth of each cell (Fig. 2). Analysis was by a ruler rather than a FACS. Growth appeared to be unaffected by the coating procedures, as we used short centrifugation times and worked in a warm room at 36°C to avoid releasing the mutant block. Pulse times below 30 min were impracticable because the overlap mentioned above reduced contrast between the bands.

Cells which were small initially grew at one end and those which were large grew at both ends. The critical length was 11.5 μ m at 1 h after shift-up. Double-enders were difficult to analyse, partly because many of them had a different number of bands at each end (see, for example, Fig. 2). Figure 3 therefore shows some representative graphs of the growth rate of single-enders. Of 32 cells analysed, 14 showed stepwise rises in rate (Fig. 3a,b). The mid-point of the step was at a mean of 18.3 μ m (s.d. 2.36) and the mean increase in rate at the step was $\times 1.48$. This is in good accord with the results from films¹ where the equivalent values were 19.4 μ m and $\times 1.33$. Other cells reached a plateau either with a very early step (Fig. 3c) or with a prolonged one (Fig. 3d). Taking these cells together with those with well-defined steps, 20 cells reached a plateau in rate at a mean length of 19.4 μ m (s.d. 3.1). This is very similar to the results in Fig. 1 from the FACS analysis. Nevertheless, there was a substantial minority of cells (14 out of 32) which did not fit these patterns and an extreme, though rare, example is shown in Fig. 3e. This variability in pattern (also found in the film results) is real and very probably greater in blocked mutants than in normal cells, but our results do emphasize the fact that analyses of whole populations can conceal what individual cells do. A final point about Fig. 3 is that three out of the five graphs show a fall in rate at the end, in contrast to Fig. 1. This is because the experiments were continued for 7 h, 2.5 h longer than in Fig. 1, and blocked mutants show a fall-off in rate after prolonged periods.

In conclusion, these two methods confirm the results from films¹ but by very different methods. The FACS analysis has also established a plateau in rate with longer cells. This was not previously known for cell length but has been shown for enzyme potential⁶ and for the rate of synthesis of protein⁷ and ribosomal RNA⁸. We believe that the methods should have wider application in other cells which have rigid cell walls and defined growing points.

This work was supported by the SERC and the Cancer

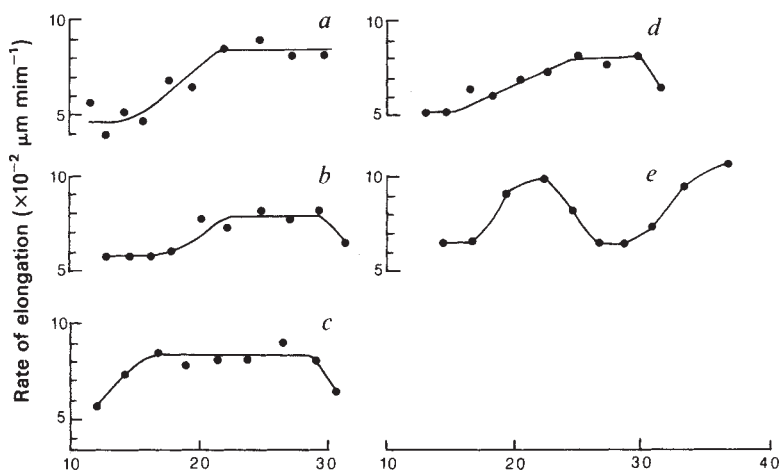


Fig. 3 Representative patterns for rate of tip growth against cell length in individual cells of *S. pombe cdc 2.33* for 7 h after shift-up. Cell growth, medium and shift-up were as for Fig. 1. Cells were taken at 1 h after shift-up; treated for 1 min in biotinylated BSL at 5 μ g ml⁻¹; washed once; 1 min in avidin at 5 μ g ml⁻¹; washed twice; resuspended in conditioned medium and incubated at 36.5°C for 30 min; then treated for 1 min in BSL-FITC at 2 μ g ml⁻¹; washed once; incubated for 30 min and the BSL-biotin-avidin coating repeated. The cycle of alternate avidin-biotin and FITC coatings was repeated every 30 min until 7 h after shift-up. The cells were finally fixed in 1% formalin and photographed in a Zeiss Photomicroscope I with a $\times 40$ objective and the exciting wavelength for FITC. The band positions were measured on photographic enlargements.

Research Campaign. We thank Drs J. Creanor, P. A. Fantes and R. A. Gray and Mr A. B. Sanderson for assistance in some experiments and for helpful discussions.

Received 21 April; accepted 4 June 1986.

1. Mitchison, J. M. & Nurse, P. *J. Cell Sci.* **75**, 357-376 (1985).
2. Yoneda, M., Ikeda, M. & Washitani, S. *Dev. Growth Differentiation* **20**, 329-336 (1978).
3. Newport, J. W. & Kirschner, M. W. *Cell* **37**, 731-742 (1984).
4. Horisberger, M., Vonlanthen, M. & Rosset, J. *Archs Mikrobiol.* **119**, 107-111 (1978).
5. Nurse, P., Thuriaux, P. & Nasmyth, K. *Molec. gen. Genet.* **146**, 167-178 (1976).
6. Benitez, T., Nurse, P. & Mitchison, J. M. *J. Cell Sci.* **46**, 399-431 (1980).
7. Creanor, J. & Mitchison, J. M. *J. Cell Sci.* **69**, 199-210 (1984).
8. Elliott, S. G. *Molec. gen. Genet.* **192**, 212-217 (1983).
9. Creanor, J. & Mitchison, J. M. *J. Cell Sci.* **58**, 263-285 (1982).

Propulsion of organelles isolated from *Acanthamoeba* along actin filaments by myosin-I

Richard J. Adams & Thomas D. Pollard

Department of Cell Biology and Anatomy, Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205, USA

Eukaryotic cells are dependent on their ability to translocate membraneous elements about the cytoplasm. In many cells long translocations of organelles are associated with microtubules¹⁻³. In other cases, such as the rapid cytoplasmic streaming in some algae, organelles appear to be propelled along actin filaments⁴. It has been assumed, but not proven, that myosin produces these movements. We have tested vesicles from another eukaryotic cell for their ability to move on the exposed actin bundles of *Nitella*⁵ as an indication that actin-based organelle movements may be a general property of cells. We found that organelles from *Acanthamoeba castellanii* can move along *Nitella* actin filaments. Here, we report two different experiments indicating that the single-headed non-polymerizable myosin isozyme myosin-I (ref. 6) is responsible for this organelle motility. First, monoclonal antibodies to myosin-I inhibit movement, but antibodies that inhibit double-headed myosin-II do not. Second, ~20% of the myosin-I in homogenates co-migrates with motile vesicles during Percoll density-gradient ultracentrifugation. This is the first indication of a role for myosin-I within the cell and supports the suggestion of Albanesi *et al.*⁷ that myosin-I moves vesicles in this way.

We saw directed motion of vesicles and aggregates of vesicles from *Acanthamoeba* on the substrate of *Nitella* actin bundles (Fig. 1). Movements were in the same direction as those of endogenous algal organelles but substantially slower (Fig. 2b),

the average rates of movement being $0.24 \pm 0.11 \mu\text{m s}^{-1}$ (mean $\pm$ s.d., $n = 50$) compared with $\geq 40 \mu\text{m s}^{-1}$ for endogenous movements⁸. The distribution of rates of movement for *Acanthamoeba* vesicles was wide (Fig. 2b) and varied from batch to batch. The vesicles in some preparations were never seen to move; the cause of this variation is unknown. Further, in any one extract, not all of the vesicles that settle onto the substrate were seen to move. We could follow moving vesicles for tens or hundreds of micrometres; the data in Fig. 2 represent the average rate measured over a distance of 20-140 μm . Rates of motion were relatively constant over these distances without the saltations or interruptions characteristic of organelle movements within live cells. Most of the movements that ceased in the field of view appeared to result either from damage in the substrate or from steric interference. We are confident that the movements of *Acanthamoeba* vesicles is independent of those of *Nitella* organelles for several reasons: no endogenous organelles are seen to move with this speed and all that show any signs of movement are lost from the preparation soon after dissection. (Experiments were not performed on dissected *Nitella* preparations while endogenous organelles were still moving.) To check this, two preparations of dissected *Nitella* were incubated in 5 mM *N*-ethylmaleimide for 5 min to inhibit the activity of the endogenous myosin species⁹. After treatment with 7 mM dithiothreitol to neutralize the *N*-ethylmaleimide, no endogenous organelles moved but organelles in *Acanthamoeba* extracts moved normally. The enzyme responsible for their motion would, therefore, appear to be associated with the vesicles themselves.

Acanthamoeba contains two distinct species of myosin^{6,10,11} which are candidates for the motor driving of these vesicles along the actin bundles of the *Nitella* cortex. The two species differ, both physically and enzymatically: myosin-I is enzymatically highly active⁶ but, unlike *Acanthamoeba* myosin-II and other known myosins, it consists of only one heavy chain which lacks the long α -helical tail required for polymerization into thick filaments^{6,7}. Like most other myosins, myosin-II has two heads and a tail and can form bipolar filaments.

We used monoclonal antibodies to show that myosin-I but not myosin-II is required for organelle movements along actin filaments. Antibody M1.5 inhibits stoichiometrically the actin-activated ATPase of myosin-I (ref. 12) and antibodies M2.3, M2.10 and M2.26 all stoichiometrically inhibit the actin-activated ATPase of purified myosin-II as well as the contraction of actin gels in crude cytoplasmic extracts¹³. These four antibodies react with the heavy chains of either myosin-I (M1.5) or myosin-II (M2.3, M2.10 and M2.26) among the various proteins in cellular extracts^{12,14}. M1.5 and other monoclonal antibodies to myosin-I also bind to nuclear proteins¹². In blind experiments antibody M1.5 completely stopped all organelle movement on

Table 1 Antibody inhibition of movements

Experimental preparation	Antigen	Antibody: myosin ratio	Presence of movement
Untreated extract			+
Untreated extract on substrate pretreated with 5 mM <i>N</i> -ethylmaleimide then dithiothreitol			+
Extract treated with antibodies*:			-
M2.3	Myosin-II	10:1	+
M2.10	Myosin-II	20:1	+
M2.26	Myosin-II	20:1	+
M1.5	Myosin-I	20:1	-
		2:1†	+/-
		0.2:1	+
		0.02:1	+
Alice‡	Chicken muscle myosin	-	+

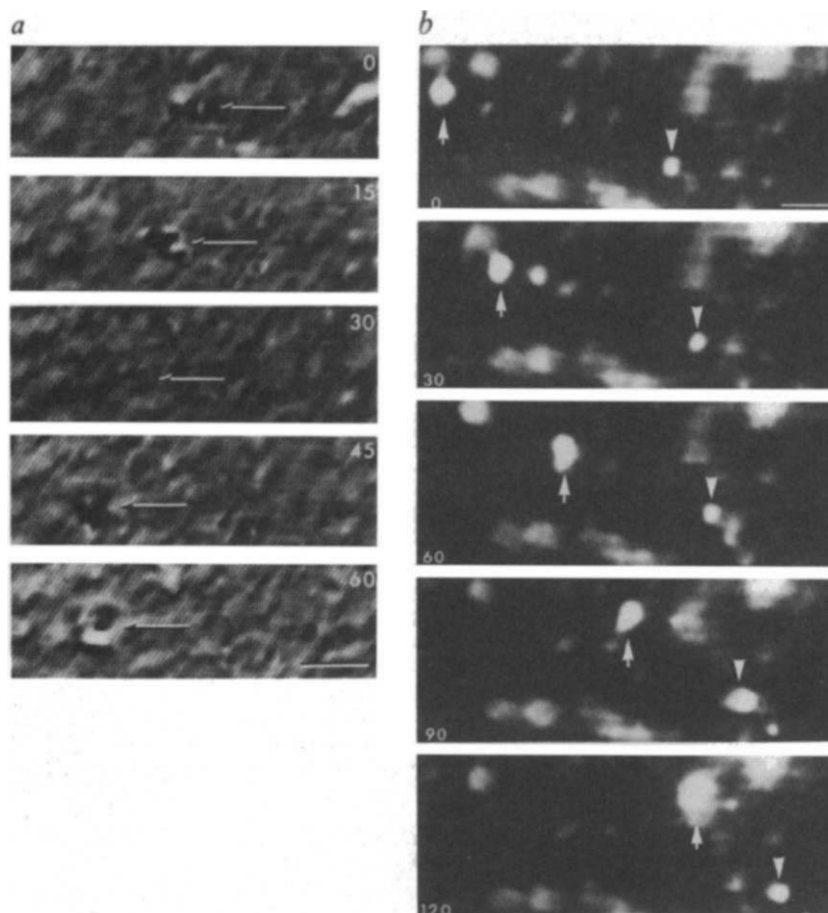
* Results of blind experiments in which the experimenter did not know the treatment of the organelle preparations.

† Results of two sets of blind experiments carried out at this ratio of antibody to antigen showed inhibition in one but not the other experiment.

‡ Monoclonal antibody raised against chicken muscle myosin; does not cross-react with *Acanthamoeba* myosins.

Fig. 1 *a*, Organelles (arrows) moving on a substrate of *Nitella* actin bundles, photographed at 15-s intervals. The image-processing procedure (see below) causes the image of the organelles to change from frame to frame. Photographs prepared from moving organelles recorded on videotape. Scale bar, 10 μm . Actin bundles oriented horizontally. *b*, Phagocytic vesicles rendered visible by feeding the cells fluorescent beads before homogenization. Fluorescent organelles (arrows) are seen against the autofluorescence of the *Nitella* chloroplasts. Photographs taken at 30-s intervals from an unprocessed video-recording. Scale bar, 10 μm . Actin bundles oriented horizontally.

Methods. *Acanthamoeba castellanii* were grown in liquid culture as described previously⁶. Organelles were prepared freshly each day starting from approximately 0.2 g wet weight of compacted cells, washed three times in 50 mM NaCl. The cells were resuspended in 1 ml of homogenization buffer (60 mM potassium glutamate, 2 mM EGTA, 2 mM MgCl_2 , 10 mM imidazole pH 6.4, 0.1 mM benzamide, 1 mM phenylmethylsulphonyl fluoride and 4 mM ATP) and disrupted by eight strokes with a tight-fitting pestle in a 7-ml Dounce homogenizer. The homogenate was cleared of unbroken cells and large particles with a 5-s spin at 10,000 g in a bench-top Eppendorf centrifuge. The resulting supernatant was used for all the experiments. In some experiments cells were prepared with fluorescent phagocytic vesicles. Washed cells were incubated, with occasional mixing, for 15 min at room temperature in a dense suspension of rhodamine-labelled polyacrolein beads in 50 mM NaCl. (Beads of $\sim 0.5 \mu\text{m}$ diameter were prepared by the method of Margel *et al.*²³). Labelled cells



were further washed three times in cold saline and processed as before. *Nitella* was dissected as described by Sheetz and Spudich⁵ in a buffer of 20 mM KCl, 4 mM EGTA, 4 mM MgCl_2 , 10 mM sucrose, 5 mM imidazole pH 7.0 and 2 mM ATP. The suspension of organelles was mixed 1:1 with homogenization buffer containing 0.2 M sucrose and added to a region of exposed *Nitella* ectoplasm with a microcapillary. Motion of organelles on *Nitella* was observed on an upright Zeiss microscope with a $\times 40$ water-immersion, phase-contrast objective. At this magnification movements were too slow to be discerned by eye but became obvious when speeded up by time-lapse video microscopy. The microscope was fitted with Dage-MIT Newvicon video camera, the output of which was passed through a Quantex digital image processor (DS 30) and Colorado Video analogue video processor (604) before being recorded at 1/18 real-time on a National Panasonic (NV 8030) time-lapse video recorder. Detection of movement of small vesicles on the complex substrate was difficult, so the following image-processing regime was used to selectively enhance moving particles. The image processor continuously subtracted an image of the non-moving *Nitella* background from the incoming picture. Thus, only differences between the two images are seen, that is, objects that moved. This technique is complicated because the *Nitella* substrate distorts to some extent during the experiment (~ 20 min). To overcome this, the background image was itself slowly averaged using a rolling average, with the incoming image, that replaces the stored image every 128 frames. It is necessary to ensure that the rate of averaging is slower than the rate of detection of the differences (the rate of recording on the video recorder). The contrast of the difference image from the digital processor was further enhanced with the Colorado Video analog processor.

Nitella actin bundles, whereas antibody inhibitors of myosin-II (M2.3, M2.10 and M2.26) had no obvious effect on organelle movements (Fig. 2*a*). The average rate of movement was $0.23 \pm 0.09 \mu\text{m s}^{-1}$ (mean $\pm$ s.d., $n = 12$) in the presence of myosin-II antibodies. A ratio of M1.5 antigen binding sites (this antibody is an immunoglobulin M) to estimated myosin-I content of the extract (1.3 μmol per kg of wet cells⁶) of 20:1 reproducibly inhibited motion whereas a ratio of 2:1 inhibited motion in one but not a second of two sets of experiments. Ratios of 0.2:1 or 0.02:1 did not inhibit movement (Table 1). These experiments provide evidence that the movement of the organelles is driven by the smaller myosin species, myosin-I. The rate of motion of these organelles is ~ 10 times higher than that recently reported for plastic beads coated with myosin-I on a *Nitella* preparation⁷.

The composition of the homogenate used for these experiments is highly heterogeneous so we do not yet know which organelles are moving. But by pre-labelling phagocytic vesicles of the cells with fluorescent beads we could show that these were among the organelles that move in our experiments (Fig. 1*b*). This could be a subset of the total moving population.

We used cell fractionation to test whether myosin-I is physically associated with the motile vesicles (Fig. 3). When the crude extracts used in the motility experiments were centrifuged on a Percoll gradient, organelles sedimented into the density gradient and were separated from the soluble components at the top of the tube. Organelles that came to equilibrium at a density of $1.08\text{--}1.13 \text{ g ml}^{-1}$ were motile on *Nitella* actin bundles. This fraction was also greatly enriched in myosin-I relative to myosin-II, which remains essentially in the supernatant. Approximately 20% of the myosin-I in the extract sedimented with these motile organelles. Tests with specific monoclonal antibodies¹⁰ showed that both myosin-IA and -IB are associated with organelles. Therefore, there is a stable interaction of myosin-IA and -IB with organelles that persists through centrifugation, is insensitive to ATP and is competent to move organelles along actin bundles.

These experiments show that an actomyosin system can provide the force for some membrane movements within the cell, as has been speculated based on experiments with myosin-coated beads⁵⁻⁷. The heterogeneity of the organelles in our preparation and the fact that we used a cell-free system do not

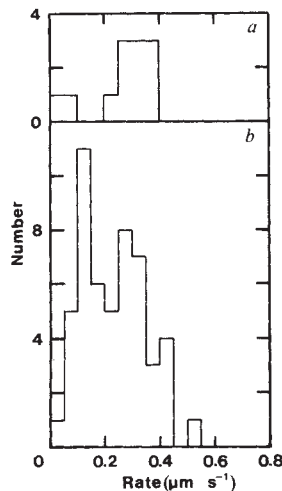


Fig. 2 Distribution of velocities of *Acanthamoeba* vesicles on a *Nitella* actin bundle substrate. *a*, Movement of 12 organelles in the presence of antibodies to myosin-II; *b*, accumulated results from 50 organelles measured in control preparations.

Methods. The net speed of movement for a vesicle was measured from the replay of a videotape recording of an experiment. The positions at the beginning and end of a timed motion were marked on the video monitor. The screen coordinates of these two points were read using a Colorado Video analyser (321) and the corresponding displacement calculated with account being made of lateral distortions introduced by the video camera. Speeds were recorded for displacements of 20–140 μm .

yet permit us to identify directly the nature or role of these movements in the cell. Immunofluorescent localization of myosin-I in the cell^{12,15} reveals a cortical concentration which suggests an involvement of myosin-I-powered movements in endocytosis or surface movements in addition to organelle movements. Such processes may be important in other cells since myosin-I has now been identified in *Dictyostelium*¹⁶ and a

similar ATPase has been identified as the link between the actin filaments and the membrane in microvilli¹⁷. Previous results which suggested a possible role for actin filaments in organelle movements in other cell types^{1,18–22} may also be mediated by proteins like myosin-I.

This work was supported by NIH research grants GM 26338 and GM 26132 (to T.D.P.) and a Postdoctoral Fellowship from the Muscular Dystrophy Association (to R.J.A.).

Received 10 March; accepted 2 June 1986.

- Schliwa, M. *Cell Muscle Motil.* **5**, 1–82 (1984).
- Schnapp, B. J., Vale, R. D., Sheetz, M. P. & Reese, T. S. *Cell* **40**, 455–462 (1985).
- Allen, R. D. *et al. J. Cell Biol.* **100**, 1736–1752 (1985).
- Kachar, B. *Science* **227**, 1355–1357 (1985).
- Sheetz, M. P. & Spudich, J. A. *Nature* **303**, 31–35 (1983).
- Pollard, T. D. & Korn, E. D. *J. biol. Chem.* **248**, 4682–4690 (1973).
- Albanesi, J. P. *et al. J. biol. Chem.* **260**, 8649–8652 (1985).
- Kamiya, N. & Kuroda, K. *Bot. Mag.* **69**, 544–554 (1956).
- Chen, J. C. W. & Kamiya, N. *Cell Struct. Funct.* **1**, 1–9 (1975).
- Maruta, H. & Korn, E. D. *J. biol. Chem.* **252**, 6501–6509 (1977).
- Pollard, T. D., Stafford, W. F. & Porter, M. E. *J. biol. Chem.* **253**, 4798–4808 (1978).
- Hagen, S. C., Kiehart, D. P., Kaiser, D. A. & Pollard, T. D. *J. Cell Biol.* (in the press).
- Kiehart, D. P. & Pollard, T. D. *J. Cell Biol.* **99**, 1024–1033 (1984).
- Kiehart, D. P., Kaiser, D. & Pollard, T. P. *J. Cell Biol.* **99**, 1002–1014 (1984).
- Gadasi, H. & Korn, E. D. *Nature* **286**, 452–456 (1980).
- Cote, G. P., Albanesi, J. P., Ueno, T., Hammer, J. A. & Korn, E. D. *J. biol. Chem.* **260**, 4543–4546 (1985).
- Collins, J. H. & Borysenko, C. W. *J. biol. Chem.* **259**, 14128–14135 (1984).
- Edds, K. T. *J. Cell Biol.* **66**, 145–155 (1975).
- Bradley, T. J. & Satir, P. *J. supramolec. Struct.* **12**, 165–175 (1979).
- Isenberg, G., Schubert, P. & Kreutzberg, G. W. *Brain Res.* **194**, 588–593 (1980).
- Goldberg, D. J., Harris, D. A., Lubit, B. W. & Schwartz, J. H. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7448–7452 (1980).
- Brady, S. T., Lasek, R. J., Allen, R. D., Yin, H. L. & Stossel, T. P. *Nature* **310**, 56–58 (1984).
- Margel, S., Beitel, U. & Ofarim, M. *J. Cell Sci.* **56**, 157–175 (1982).
- Laemmli, U. K. *Nature* **229**, 680–685 (1970).
- Towbin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350–4354 (1979).

Lateral proton conduction at lipid–water interfaces and its implications for the chemiosmotic-coupling hypothesis

Michel Prats, Justin Teissie
& Jean-François Tocanne*

Centre de Recherche de Biochimie et de Génétique Cellulaires du CNRS, 118 Route de Narbonne, 31062 Toulouse Cedex, France

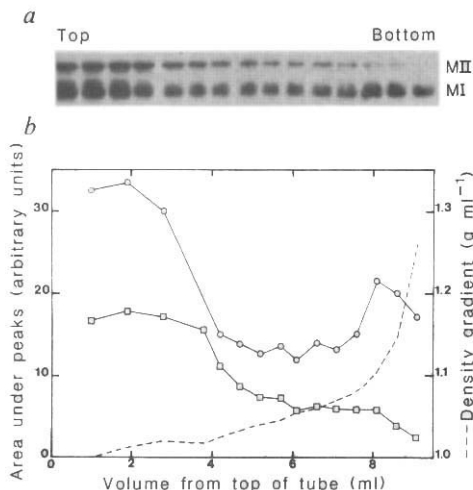


Fig. 3 *Acanthamoeba* cytoplasmic extract (0.5 ml) was layered onto 9.0 ml of 1.08 g ml^{-1} Percoll in homogenization buffer and centrifuged at 27,000 r.p.m. in a Beckman Ti 50 rotor for 15 min. The gradient was unloaded from the bottom of the tube; the densities of Percoll in the collected fractions were determined from refractive indices. Aliquots of these fractions were electrophoresed on 7.5% gels²⁴, transferred to nitrocellulose²⁵ and radioimmunologically stained with antibodies to myosin-I and -II (*a*). The autoradiogram shown in *a* was scanned using a soft-scanning laser densitometer to show the distribution of the antigens in the fractions collected from the gradient. MI, myosin-I; MII, myosin-II. *b*, Relative areas scanned under the peaks of myosin-I and -II were plotted for each fraction through the gradient. Circles, myosin-I; squares, myosin-II.

The driving force in energy-transducing membranes is now recognized to be linked to a flux of protons between membrane-bound donors and acceptors. The question of whether the proton pathway is delocalized within the two bulk aqueous phases on each side of the membrane (delocalized chemiosmotic theory¹), or is localized at its surface (semi-localized hypothesis²) still remains open to discussion^{3,4}. Using an original fluorescence monolayer technique, we have recently been able to provide direct experimental evidence that a phospholipid–water interface can act as an efficient proton conductor^{5,6}. This observation strongly supports the semi-localized hypothesis². In the present study, comparisons between surface potential and fluorescence measurements in monolayers provide additional evidence of a facilitated proton conduction along phospholipid–water interfaces. They suggest the existence of an induced steep surface pH gradient from a more acidic surface towards the bulk. They also show that lateral proton transfer along the surface of a biological membrane would alter the surface potential of that membrane.

The surface potential ΔV which is measured for lipids in monolayers is known to originate primarily from the polar headgroups and to depend on their ionization state⁷. Therefore, since the lipid is the probe itself, this technique affords an elegant way of observing the lateral movements of protons and of

* To whom correspondence should be addressed.

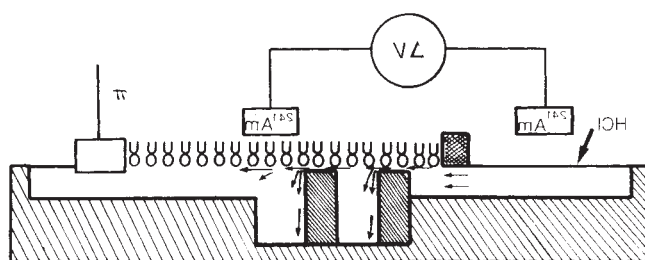


Fig. 1 Schematic drawing of the experimental set-up. The monolayer experiments were carried out with the same trough and using the same experimental conditions as previously described for fluorescence experiments^{5,6}. The trough comprised one 'injection' and one 'observation' compartment separated by two glass barriers. Continuity of the subphase occurred via a very thin water layer, about 1 mm in depth, between the water surface and the top of the barriers. π was measured by means of a platinum plate connected to a force transducer of our own manufacture. Surface potential was measured by two americium electrodes¹¹. One (reference) was located above the lipid-free region of the 'injection' compartment. The other (measurement) faced the lipid film. The 'standard' conditions^{5,6} we used were: the 'measurement' electrode was centred at a distance of 4.3 cm from the first glass barrier; the area of the monolayer in direct contact with the acidic subphase in the 'injection' compartment was 12.6 cm²; protons (150 μ l 3 M HCl) were injected under stirring (Teflon bar, 90 r.p.m.). 'Standard' conditions of subphase (1 mM phosphate, pH 6.8) and film surface pressure (25 mN m⁻¹) were used. Arrows indicate the various diffusion pathways of protons which can move along the lipid/water interface and which can also diffuse through the bulk of the water phase by a combination of diffusion and convection movements.

determining to what extent the polar headgroups are involved in the proton conduction process.

Three phospholipids were examined: phosphatidylethanolamine (PE) from *Escherichia coli* (Sigma), phosphatidylserine (PS) from bovine brain (Sigma) and dilauroylphosphatidylglycerol (PG) of synthetic origin⁸. As shown in Fig. 1, the trough, the method for generating the flux of protons and the experimental conditions used were identical to those previously described for the fluorescence experiments^{5,6}.

Typical ΔV recordings with time are shown in Fig. 2. Less than 3 min after proton injection, a large increase in ΔV was observed for each lipid. A positive change in ΔV is expected if protons concentrate at the lipid-water interface⁷. From these curves, two parameters were determined: (1) $TH_{\Delta V}^+$, the time which is necessary to detect a change in the surface potential and which relies on the rate of proton transfer along the lipid/water interface^{5,6}, (2) $\Delta\Delta V$, the change in surface potential which, like the change in surface fluorescence ΔF (refs 5, 6), accounts for the variation in proton concentration at the interface. $TH_{\Delta V}^+$ and $\Delta\Delta V$ values measured for each lipid are shown in Table 1.

Fluorescence investigation of proton conduction by phosphatidylserine and phosphatidylglycerol was carried out exactly as previously described for phosphatidylethanolamine^{5,6}. 'Standard' conditions of subphase (1 mM phosphate, pH 6.8) and

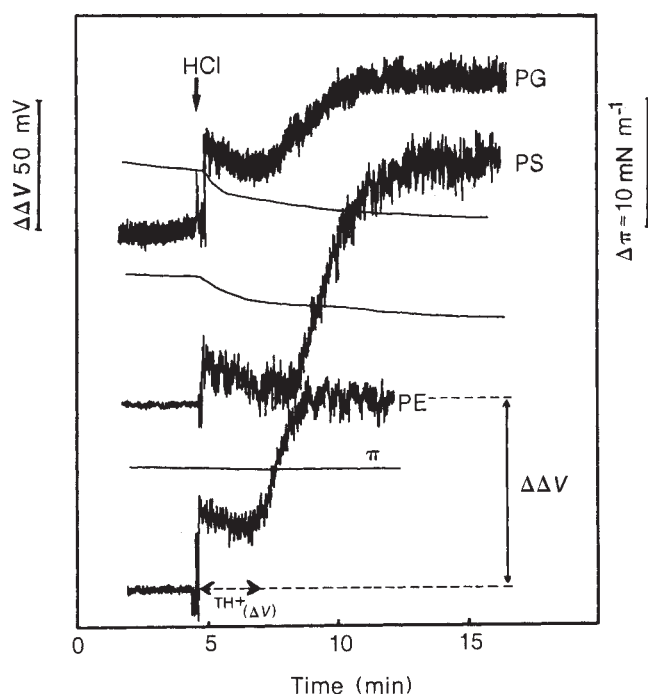


Fig. 2 Kinetics of the surface potential and surface pressure changes after an acid jump. Phosphatidylglycerol (PG), phosphatidylserine (PS) and phosphatidylethanolamine (PE) were spread at a surface pressure of 25 mN m⁻¹ by stepwise deposition of the lipids onto the subphase. The left part of the curves refers to the surface potential which was initially measured after spreading the lipids on the 1 mM phosphate subphase at pH 6.8. Addition of HCl gave rise to an immediate drop of the pH from 6.8 to 2.2 in the 'injection' compartment. This resulted in a rise of the surface potential due to polarization of the 'reference' electrode. Control experiments carried out in the absence of film showed that this rise in ΔV rapidly vanished with time. After about 5 min, the ΔV returned to its initial value. Changes in ΔV resulting from proton translocation along the various lipid monolayers were therefore determined from the surface potential measured before the addition of HCl, which can be considered as a 'reference' potential. $TH_{\Delta V}^+$ is the delay between H⁺ injection and the beginning of the change in ΔV . $\Delta\Delta V$ is the change in surface potential which is measured at equilibrium, when a plateau is again observed. The thin continuous curves are the recorder traces of the surface pressure π .

surface pressure (25 mN m⁻¹) were used. The TH_F^+ and ΔF values which were measured for the two lipids are shown in Table 1, together with those already reported for PE^{5,6}. As expected, the two techniques lead to similar TH^+ (around 2 min) for each lipid. From these values, PS would be a less efficient proton translocator than PE or PG, which exhibit the same behaviour.

The change in 'interfacial' pH which originates from proton conduction cannot be determined directly from the ΔF or $\Delta\Delta V$ values of Table 1. Nevertheless, and as a first approximation, this change in 'interfacial' pH can be related to the change in subphase pH for which the same ΔF and $\Delta\Delta V$ values would

Table 1 Changes in surface potential, surface fluorescence and pH for various phospholipids

Phospholipids	Surface potential				Fluorescence			
	$TH_{\Delta V}^+$ (s)	$\Delta\Delta V$ (mV)	pH	ΔpH^*	TH_F^+ (s)	ΔF (%)	pH	ΔpH^*
PE	140 $\pm$ 20	61 $\pm$ 10	2.6 $\pm$ 0.2	4.2 $\pm$ 0.2	120 $\pm$ 5†	78 $\pm$ 18†	5.0 $\pm$ 0.8	1.8 $\pm$ 0.8
PG	130 $\pm$ 10	40 $\pm$ 20	3.1 $\pm$ 0.4	3.7 $\pm$ 0.4	120 $\pm$ 10	65 $\pm$ 5	5.5 $\pm$ 0.2	1.3 $\pm$ 0.2
PS	170 $\pm$ 15	75 $\pm$ 25	3.7 $\pm$ 0.5	3.1 $\pm$ 0.5	160 $\pm$ 5	41 $\pm$ 6	6.0 $\pm$ 0.2	0.8 $\pm$ 0.2

TH^+ , changes in surface potential $\Delta\Delta V$ and in surface fluorescence ΔF and measured equivalent bulk-phase pH and ΔpH values for films of phosphatidylethanolamine (PE), phosphatidylglycerol (PG) and phosphatidylserine (PS) (average of three to seven determinations)

* ΔpH is defined as the difference between the initial subphase pH (6.8) and the measured equivalent bulk-phase pH.

† From ref. 5.

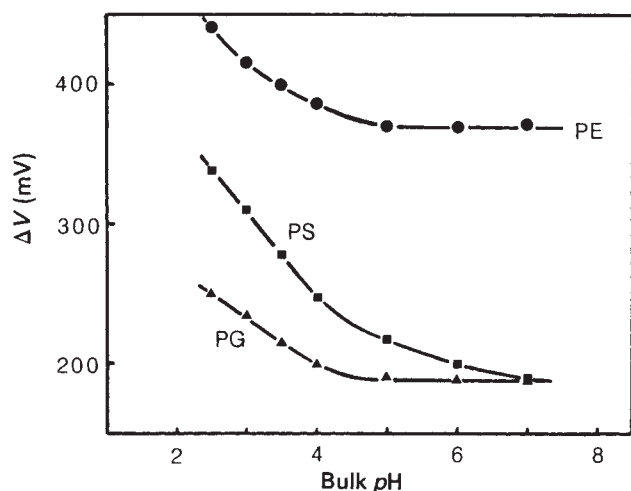


Fig. 3 Surface potential versus subphase pH titration curves. Phosphatidylethanolamine (PE), phosphatidylserine (PS) and phosphatidylglycerol (PG) were spread at a constant molecular area of 0.54 nm^2 , 0.66 nm^2 and 0.56 nm^2 respectively by stepwise deposition of the lipids onto various 1 mM phosphate subphases at the desired pH. For each lipid, these molecular areas correspond to a film surface pressure of 25 mN m^{-1} at pH 6.8. Data presented are the average of five determinations.

be measured at equilibrium in surface fluorescence and surface potential against subphase pH titration curves, by reference to the initial subphase pH of 6.8. In the following, the change in 'interfacial' pH approximated in this way will be referred to as the 'equivalent bulk-phase' ΔpH .

Surface potential/subphase pH titration curves of PE, PS and PG are shown in Fig. 3. The apparent pK measured for our fluorescent pH indicator fluorescein thiocarbamide phosphatidylethanolamine (F-PE) in the presence of PG and PS (titration curves not shown) was not significantly different from that already determined for PE⁵.

The 'equivalent bulk-phase' pH and ΔpH values determined are shown in Table 1. For each lipid, these values indicate a monolayer in contact with an interfacial water phase more acidic than the bulk water phase. The equivalent bulk-phase ΔpH values found by fluorescence were always lower than those obtained through surface potential measurements, by about 2.5 pH units. Both approaches lead to the same decreasing order of proton conduction efficiency $\text{PE} > \text{PG} > \text{PS}$. This order agrees with the observed increase in TH^+ . Furthermore, control experiments showed that the aqueous phase which was sucked from underneath the lipid film (about 2–3 mm from the surface), just after protons were detected by fluorescence or ΔV measurements, was still at the initial pH of 6.8.

For each lipid, the observed differences in 'equivalent bulk-phase' ΔpH can be accounted for by a steep pH gradient between a more acidic interface plane and the bulk of the water phase, as proposed in Fig. 4 for PE. Indeed, the surface potential senses the ionic conditions that prevail in the interface plane, where the polar headgroups are located. Fluorescein when attached to PE does not penetrate the lipid film but remains in solution in the water phase, hanging perpendicular to the interface plane⁵. The ionizable groups of this chromophore are difficult to locate accurately. From calculations based on molecular models, fluorescein should probe a thin layer of the interphase water, at a distance between 0.5 and 1.5 nm from the interface plane. A linear extrapolation of these ΔpH values over increasing distances into the bulk aqueous phase leads to the conclusion that the equivalent bulk-phase ΔpH would again be zero at a distance of about 1–3 nm from the interface plane.

Three main conclusions are to be drawn from this study. First, our fluorescence and ΔV data provide experimental evidence that proton conduction facilitated by phospholipids could occur

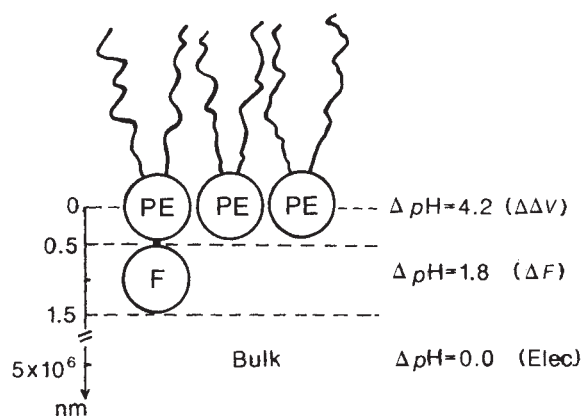


Fig. 4 Diagrammatic representation of the 'interfacial' pH gradient which is postulated to exist in the vicinity of a proton-conducting lipid film. The indicated ΔpH values refer to the 'equivalent bulk-phase' ΔpH determined in Table 1 for phosphatidylethanolamine (PE) from the corresponding surface potential and surface fluorescence data. 4.2 is the ΔpH value sensed by ΔV measurements, in the interface plane. 1.8 is the ΔpH value monitored by the fluorescent pH indicator fluorescein thiocarbamide phosphatidylethanolamine (F-PE) which probes the interphase between 0.5 and 1.5 nm from the interface plane. A ΔpH of 0 shows that the pH measured by means of a pH electrode for a few millilitres of the subphase sucked up at a distance of about 5 mm from the water surface, just underneath the ΔV 'measurement' electrode and after an increase in ΔV due to proton conduction had been observed, kept the initial value of 6.8.

at the surface of energy-transducing membranes. This provides strong support for the semi-localized chemiosmotic hypothesis². Second, in our experimental conditions, large increases in surface potential were observed for each lipid tested. This means that in the case of a localized proton pathway, changes in the surface potential of energy-transducing membranes are expected. This could have consequences for certain membrane functions⁹. Third, the comparison between fluorescence and ΔV data strongly suggests that proton conduction is accompanied by a steep pH gradient from the membrane surface towards the bulk aqueous phase, the highest proton concentration being found in the interface plane. This means that protons are picked up and transported in this interface plane, where the polar headgroups of the lipids are located. In addition, the observed differences between each lipid in TH^+ and ΔpH values (Table 1) indicate that proton conduction depends on the nature of the lipid polar headgroups, as has been postulated^{5,6,10}.

It is finally worth emphasizing that the occurrence of such an interfacial pH gradient at the surface of an energy-transducing membrane might imply that the transmembrane bulk-to-bulk water phase ΔpH which is usually measured for the determination of the protonmotive force Δp , actually does not account for the real number of protons which are exchanged in the course of energy transduction.

We thank Professor Winterton for re-reading the manuscript.

Received 25 February; accepted 4 June 1986.

- Boyer, P. D., Chance, B., Ernster, L., Mitchell, P., Racker, E. & Slater, E. A. *Rev. Biochem.* **46**, 955–1026 (1977).
- Kell, D. B. *Biochim. biophys. Acta* **549**, 55–99 (1979).
- Ferguson, S. J. *Biochim. biophys. Acta* **811**, 47–95 (1985).
- Haraux, F. *Physiol. vég.* **23**, 397–410 (1985).
- Teissié, J., Prats, M., Soucaille, P. & Tocanne, J. F. *Proc. natn. Acad. Sci. U.S.A.* **82**, 3217–3221 (1985).
- Prats, M., Tocanne, J. F. & Teissié, J. *Eur. J. Biochem.* **149**, 663–668 (1985).
- Lakhdar-Ghazal, F., Tichadou, J. L. & Tocanne, J. F. *Eur. J. Biochem.* **134**, 531–537 (1983).
- Tocanne, J. F., Verheij, H. M., op den Kemp, J. A. F. & van Deenen, L. L. M. *Chem. Phys. Lipids* **13**, 389–403 (1974).
- Wojtczak, L. & Nalecz, M. J. *Eur. J. Biochem.* **94**, 99–107 (1979).
- Haines, T. H. *Proc. natn. Acad. Sci. U.S.A.* **80**, 160–164 (1983).
- Sacré, M. M., El Mashak, E. M. & Tocanne, J. F. *Chem. Phys. Lipids* **20**, 305–318 (1977).